# Isoform-specific and signaling-dependent propagation of acute myeloid leukemia by Wilms tumor 1

Potluri, Sandeep; Assi, Salam A; Chin, Paulynn S; Coleman, Dan J L; Pickin, Anna; Moriya, Shogo; Seki, Naohiko; Heidenreich, Olaf; Cockerill, Peter N; Bonifer, Constanze

*DOI:*
10.1016/j.celrep.2021.109010



*Document Version*
Publisher's PDF, also known as Version of record

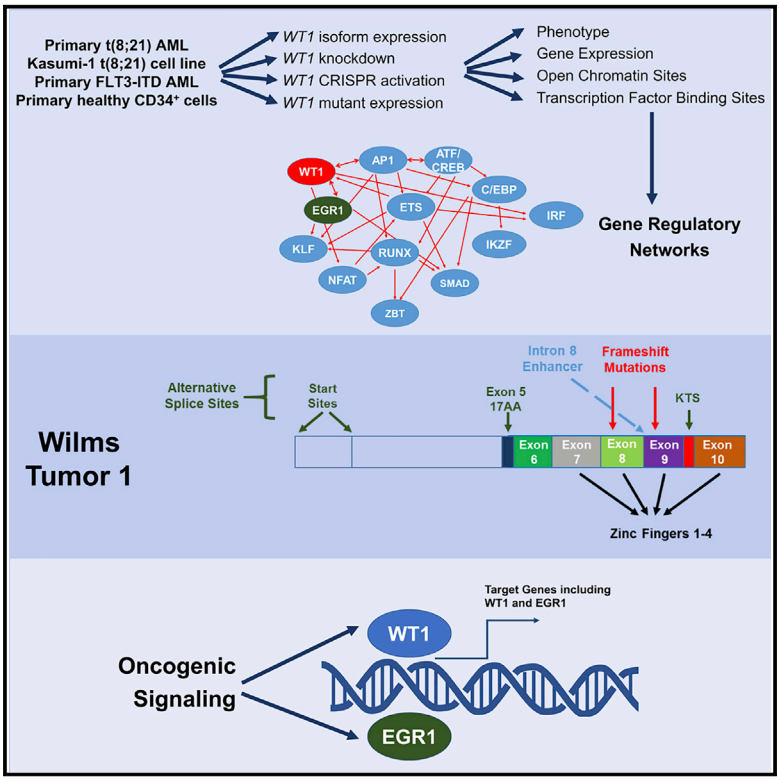
*Citation for published version (Harvard):*
Potluri, S, Assi, SA, Chin, PS, Coleman, DJL, Pickin, A, Moriya, S, Seki, N, Heidenreich, O, Cockerill, PN & Bonifer, C 2021, 'Isoform-specific and signaling-dependent propagation of acute myeloid leukemia by Wilms tumor 1', *Cell Reports*, vol. 35, no. 3, 109010. https://doi.org/10.1016/j.celrep.2021.109010

Cell Reports

Article

# Isoform-specific and signaling-dependent propagation of acute myeloid leukemia by Wilms tumor 1

## Graphical abstract




## Highlights

- WT1 is an important factor for AML maintenance in multiple subtypes
- WT1 acts in an isoform- and AML-subtype-specific fashion
- WT1 binding to chromatin is in balance with early growth response factors
- Upregulation of expression is part of an interlinked oncogenic signaling network


## Authors

Sandeep Potluri, Salam A. Assi, Paulynn S. Chin, ..., Olaf Heidenreich, Peter N. Cockerill, Constanze Bonifer

## Correspondence

s.potluri@bham.ac.uk (S.P.), c.bonifer@bham.ac.uk (C.B.)



## In brief

Potluri et al. demonstrate that in acute myeloid leukemia (AML), WT1 alters growth in an isoform-specific manner. It is regulated by oncogenic signaling and other oncogenes, plays a wide-spread role in AML biology, and is part of essential gene regulatory networks that maintain multiple leukemic subtypes.


Potluri et al., 2021, Cell Reports *35*, 109010
April 20, 2021 
https://doi.org/10.1016/j.celrep.2021.109010

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

# Isoform-specific and signaling-dependent propagation of acute myeloid leukemia by Wilms tumor 1

Sandeep Potluri,[1,*] Salam A. Assi,[1,7] Paulynn S. Chin,[1,7] Dan J.L. Coleman,[1] Anna Pickin,[1] Shogo Moriya,[2] Naohiko Seki,[3] Olaf Heidenreich,[4,5,6] Peter N. Cockerill,[1,6] and Constanze Bonifer[1,8,*]

[1]Institute of Cancer and Genomic Sciences, College of Medicine and Dentistry, University of Birmingham, Edgbaston, Birmingham B152TT, UK

[2]Department of Biochemistry and Genetics, Chiba University Graduate School of Medicine, Chiba, Japan

[3]Department of Functional Genomics, Chiba University Graduate School of Medicine, Chiba 260-8670, Japan

[4]Wolfson Childhood Cancer Research Centre, Translational and Clinical Research Institute, Newcastle University, Herschel Building, Level 6, Brewery Lane, Newcastle upon Tyne NE1 7RU, UK

[5]Prinses Máxima Centrum for Pediatric Oncology, Postbus 113, 3720 AC Bilthoven, Heidelberglaan 25, 3584CS Utrecht, the Netherlands

[6]Senior author

[7]These authors contributed equally

[8]Lead contact

*Correspondence: s.potluri@bham.ac.uk (S.P.), c.bonifer@bham.ac.uk (C.B.)

https://doi.org/10.1016/j.celrep.2021.109010

## SUMMARY

Acute myeloid leukemia (AML) is caused by recurrent mutations in members of the gene regulatory and signaling machinery that control hematopoietic progenitor cell growth and differentiation. Here, we show that the transcription factor WT1 forms a major node in the rewired mutation-specific gene regulatory networks of multiple AML subtypes. *WT1* is frequently either mutated or upregulated in AML, and its expression is predictive for relapse. The WT1 protein exists as multiple isoforms. For two main AML subtypes, we demonstrate that these isoforms exhibit differential patterns of binding and support contrasting biological activities, including enhanced proliferation. We also show that *WT1* responds to oncogenic signaling and is part of a signaling-responsive transcription factor hub that controls AML growth. WT1 therefore plays a central and widespread role in AML biology.

## INTRODUCTION

Acute myeloid leukemia (AML) is a hematological malignancy that occurs as a consequence of genetic alterations (Papaemmanuil et al., 2016). An individual AML has an average of five recurrent mutations plus a variable number of less recurrent mutations (Ley et al., 2013). Two to three of these mutations tend to be true interdependent drivers of leukemogenesis, with other mutations representing interchangeable co-operating mutations or random events (Welch et al., 2012). Of these, the majority of AMLs carry mutations conferring a block in hematopoietic differentiation and enhanced self-renewal, typically in genes encoding transcription factors (TFs) or epigenetic regulators plus at least one that constitutively activates intracellular signaling independent of growth factors. Signaling mutations tend to be acquired as secondary mutations and allow for enhanced cell proliferation and survival. The most common signaling mutations in AML are internal tandem duplications (ITDs) in the transmembrane domain of the receptor FLT3 (FLT3-ITDs) that cause its constitutive activation (Hayakawa et al., 2000). The most common TF mutations in AML are in the RUNX1/CBFβ genes encoding for Core Binding Factor, in particular the recurrent t(8;21) translocation that fuses the DNA-binding domain of RUNX1 to almost the entire ETO (RUNX1T1) repressor protein, generating the oncogenic fusion protein RUNX1-ETO (Erickson et al., 1992).

Normal hematopoietic cell differentiation originates from hematopoietic stem cells (HSCs) and involves the differential expression of lineage-determining TFs. Gene regulatory networks (GRNs) consist of constitutive and signaling responsive TFs and the genes that they bind to and regulate, which themselves may encode TFs. The binding of TFs to developmental-stage-specific target genes therefore drives shifts in GRNs from one progenitor cell type to the next until terminal differentiation is reached (Goode et al., 2016). These finely balanced GRN dynamics are disturbed once one of their essential components is mutated and terminal differentiation cannot be properly executed.

To investigate the relationship between AML subtypes and leukemic GRNs, we recently conducted a multi-omics analysis of seven AML subtypes defined by mutations in the genes encoding the TFs RUNX1/CBFβ and C/EBPα, the growth factor receptor mutation FLT3-ITD, or Nucleophosmin 1 (NPM1) which often co-occurs with FLT3 mutations (Assi et al., 2019b). Our analyses demonstrated that depending on the type of mutation,

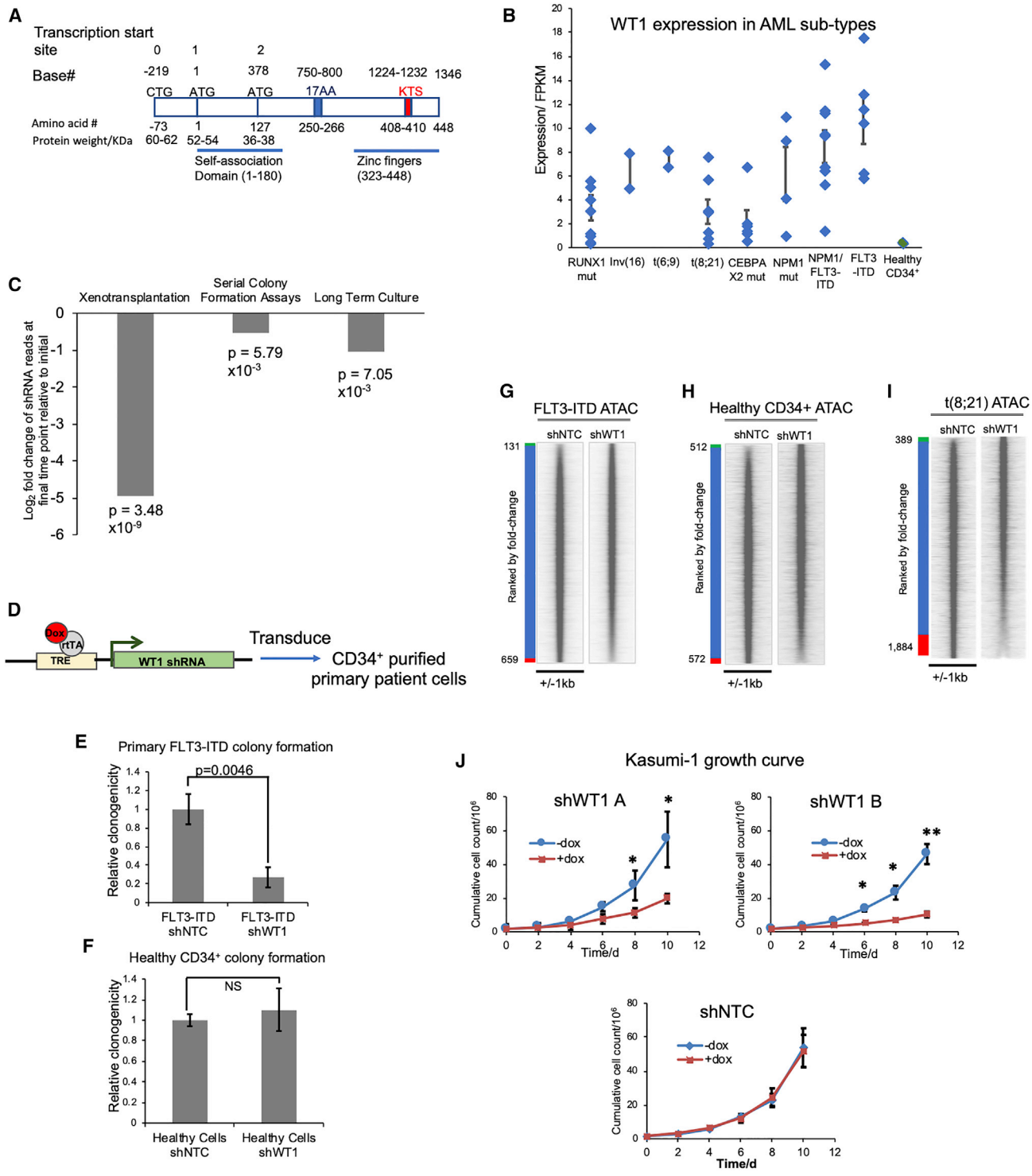



**Figure 1. WT1 is essential for leukemic maintenance**

(A) Schematic of WT1 structure and sites of alternate start sites and splice sites.

(B) Expression of *WT1* mRNA in purified primary AML leukemic blasts with different mutational subtypes (data from Assi et al., 2019b). Error bars show standard error of the mean.

(C) Results of the shRNA drop-out screen conducted by Martinez-Soria et al. (2019) for *WT1*. Log2 fold change of shRNA reads at the final time point relative to initial time point in xenotransplantation, serial colony formation assays, or long-term culture.

(D) Schematic of doxycycline-inducible shRNA construct targeting WT1 or NTC (non-targeting control).

*(legend continued on next page)*

leukemic blast cells from each AML subtype display a specific GRN with TFs forming different network nodes and being connected to other nodes as well as subtype specifically expressed target genes. These analyses showed that each AML subtype establishes its own cellular identity that differs from that of normal cells, i.e., differentiation has gone "side-ways." For selected TFs forming prominent network nodes within an AML-subtype-specific GRN, we showed that they consist of TFs/TF families that are required for the maintenance of the leukemic phenotypes of their own subtype but not others (Assi et al., 2019a; Assi et al., 2019b). Moreover, they are not essential for the growth of normal cells, highlighting an increased dependency of cancer GRNs on specific pathways that could be exploited for therapeutic purposes. Importantly, our analysis also found some TF families forming network nodes in more than one AML subtype. Such TF families are of great interest for the development of novel therapies, as they may regulate crucial common downstream target genes or may themselves be targets for therapeutic interventions. However, to be able to go along this path, the role of such factors in different AML subtypes has to be studied in detail, using both molecular and system-level approaches. One of these TF families consists of the AP-1 family of TFs that forms a prominent network node in the GRNs of all investigated AML subtypes which we showed to be required for the *in vitro* and *in vivo* growth of cells from two leukemic subtypes but not for normal cells (Assi et al., 2019b). Another such network node is formed by the $Zn^{2+}$-finger TF Wilms tumor 1 (WT1). WT1 plays an important role in development, as *Wt1* null mice die at embryonic day 13.5 due to renal and cardiac problems (Kreidberg et al., 1993). Even a deletion of *Wt1* in 6-week-old mice led to rapid death with renal failure and bone and fat loss (Chau et al., 2011).

*WT1* was first identified as a tumor suppressor gene that was commonly mutated in Wilms' tumor, a specific type of pediatric renal cancer (Call et al., 1990), and is aberrantly expressed in a variety of other cancers, such as those of breast, brain, and colon. Functional validation of its role as an oncogene in Wilms' tumor was shown by RNA interference whereby depletion caused cell cycle arrest and apoptosis (Tuna et al., 2005). WT1 biology and its impact on gene expression are highly complex, as it is expressed as different isoforms that display different splicing variants and encode proteins with different functional properties (Moriya et al., 2008). WT1 has been shown to interact with TET2 ,and its expression levels impact the epigenetic landscape (Rampal et al., 2014; Wang et al., 2015). WT1 also affects post-transcriptional processing of RNA, as it localizes and interacts with RNA splicing factors (Larsson et al., 1995) and is incorporated into functional spliceosomes (Davies et al., 1998).

*WT1* is mutated in about 10% of AML patients, conferring a poor prognosis (Inoue et al., 1994), and is overexpressed in 90% of AML (King-Underwood et al., 1996) at both the RNA and the protein level (Alanazi et al., 2020). The relapse-free survival of *WT1*-mutated AML patients is significantly impaired (Hou et al., 2010). *WT1* overexpression in AML correlates with a worse relapse-free survival and response to chemotherapy in AML (Barragán et al., 2004), and its expression has been used to predict relapse in patients in remission (Inoue et al., 1994). Based on these results, WT1 has been flagged as a therapeutic target in different cancers. Recently, T cell receptor gene therapies were developed and trialed on patients with AML targeting WT1 epitopes, which prevented relapse (Chapuis et al., 2019). However, mechanistically, little is known about how WT1 drives a more aggressive AML and why it is upregulated. Gene expression data from WT1-mutated patients suggest a perturbation of cell proliferation, but little more is known to date (Becker et al., 2010). It is unclear whether WT1 is an essential factor for leukemic maintenance in different AML subtypes, how WT1 mutant proteins alter cellular phenotypes, how it cooperates with other oncogenes, which target genes are bound by it, and whether the binding pattern and the impact on gene expression differ between isoforms.

Here, we addressed these questions at the gene-specific and system-wide level by studying the role of WT1 in two major AML subtypes, namely, t(8;21) and FLT3-ITD. We identified WT1 binding sites and show that WT1 is an essential part of the GRN and regulates the growth of both subtypes in an isoform-specific fashion. Our work also demonstrates that WT1 is differentially connected to other TFs within subtype-specific GRNs and impacts different sets of genes. Most importantly, we show that *WT1* responds to oncogenic signaling, indicating that it is downstream of a rewired signaling network that maintains an aggressive leukemic phenotype.

## RESULTS

### WT1 expression is required for the growth of two major AML subtypes

The full-length protein WT1 spans 448 amino acids and harbors multiple domains, including 4 $Zn^{2+}$-finger domains at the C-terminal end of the protein (Figure 1A; Hamilton et al., 1995). Differential promoter usage and splicing generate several isoforms of the protein with potentially different functions. Of particular importance are the production of the *WT1* −KTS and *WT1* +KTS isoforms that differ in sequences between zinc fingers three and four, which are generated by differential splicing at exon 9, resulting in either the presence or absence of three amino acids (lysine, threonine, and serine). This feature has

(E and F) Number of colonies formed per 2,000 cells seeded in methylcellulose, normalized to the short hairpin NTC (shNTC) control using primary FLT3-ITD AML (E) or primary $CD34^+$ bone marrow cells from a healthy donor transduced with doxycycline-inducible shNTC or sh*WT1* (F). n = 3 biological replicates, and error bars show standard deviation. Significance was determined using a two-tailed Student's t test.

(G–I) Density plots of ATAC-seq peaks in a 2-kb window, ranked top to bottom by relative tag count of peaks from doxycycline-inducible sh*WT1*-transduced cells relative to doxycycline-inducible shNTC-transduced cells in FLT3-ITD primary AML cells (G), primary $CD34^+$ bone marrow cells (H), or t(8;21) primary AML cells (I).

(J) Time course of cumulative cell count of Kasumi-1 cells transduced with one of two sh*WT1* or NTC plasmids, with or without induction of doxycycline. n = 3 biological replicates, and error bars show standard deviation. Significance was determined using a two-tailed Student's t test. *$p < 0.05$, **$p < 0.005$.

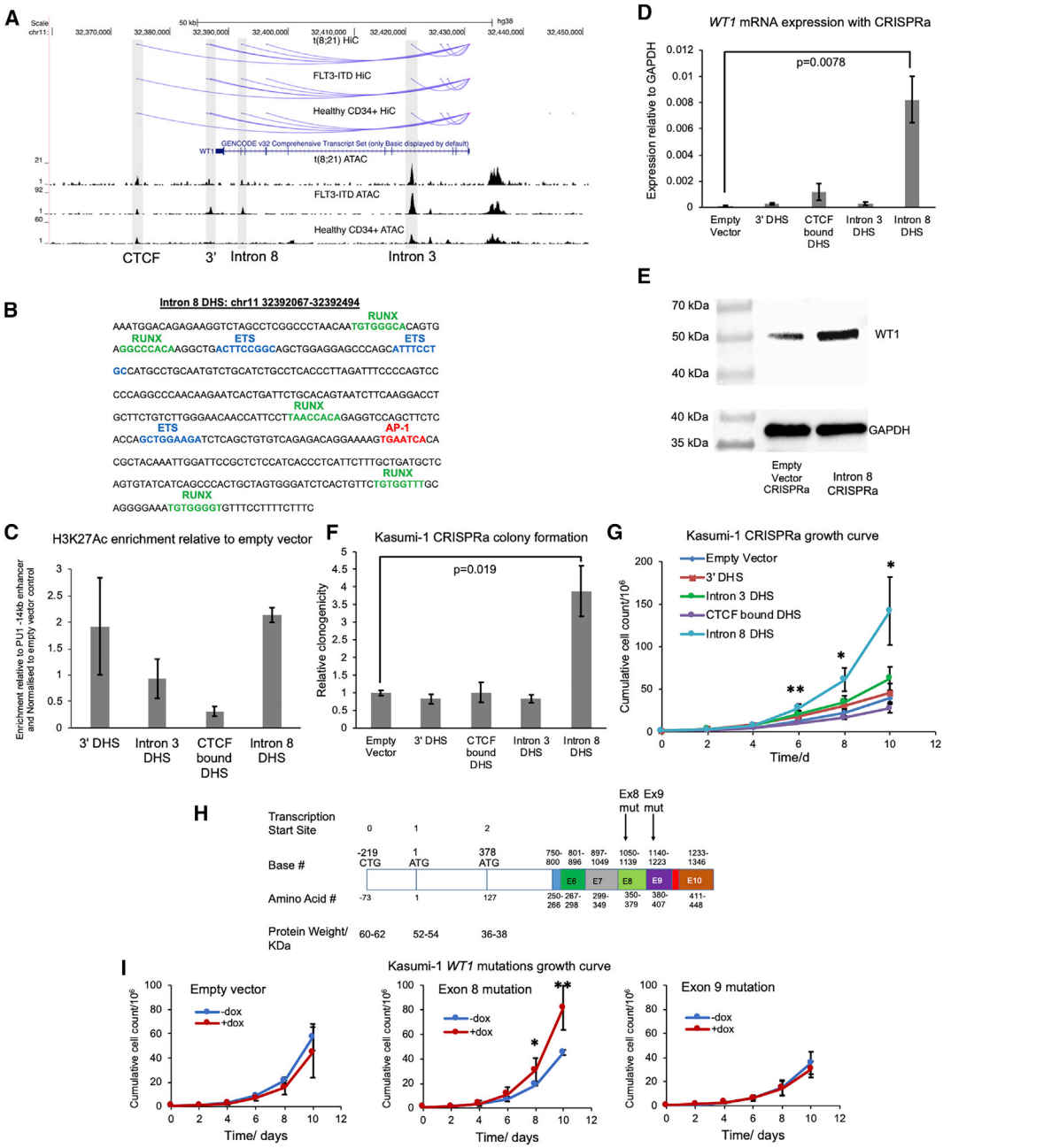



**Figure 2. *WT1* gene regulation and WT1 mutations**

(A) UCSC genome browser screenshot showing promoter capture HiC interactions and ATAC-seq peaks at the WT1 locus in primary t(8;21), primary FLT3-ITD, and healthy CD34$^{+}$ bone marrow cells. Data from Ptasinska et al. (2019) and (Assi et al., 2019b)

(B) DNA sequence at the intron 8 open chromatin site and with highlighted transcription-factor-binding motifs.

(C) H3K27ac ChIP-qPCR enrichment of Kasumi-1 cells transduced with a CRISPRa vector targeted to intron 8; levels were normalized to the empty vector control and to the PU.1 −14-kb enhancer (positive control).

(D) Relative expression levels of *WT1* mRNA normalized to *GAPDH* in Kasumi-1 cells transduced with vectors expressing CRISPRa targeted to the indicated DHSs.

(E) Western blot showing WT1 and GAPDH protein expression in Kasumi-1 cells transduced with an empty vector or with sgRNA targeted to intron 8 DHS.

(F) Colony-formating ability in methylcellulose of 2,000 Kasumi-1 cells expressing CRISPRa targeted against the indicated WT1 DHSs and normalized to the empty vector control.

(G) Time course of cumulative cell growth of Kasumi-1 cells expressing CRISPRa targeted against the indicated WT1 DHSs.

*(legend continued on next page)*

functional consequences, as it leads to different DNA-binding properties of WT1 (Bickmore et al., 1992; Laity et al., 2000). *WT1* mRNA is not expressed in purified $CD34^+$ mobilized peripheral blood stem cells (PBSCs) but is overexpressed in purified leukemic blasts of all AML subtypes that we have analyzed to date (Figure 1B) and WT1 represents a network node in multiple mutation-specific AML subtypes (Assi et al., 2019b). In our study, we focused on the role of WT1 in two AML subtypes that we have extensively characterized at the multi-omics level, namely, the t(8;21) and FLT3-ITD AML (Cauchy et al., 2015; Ptasinska et al., 2014). Both AML subtypes display distinct GRNs (Figures S1A and S1B). We have previously shown that knockdown of *WT1* expression in the context of an RNAi screen strongly affects leukemic growth of t(8;21) cells *in vitro* and after xenotransplantation *in vivo* (Martinez-Soria et al., 2019; Figure 1C). To examine whether this was also true for primary cells from AML patients, we transduced purified healthy-donor-derived $CD34^+$ stem cells and FLT3-ITD and t(8;21) AML blast cells with lentiviral vectors expressing an short hairpin RNA (shRNA) targeted against all isoforms of *WT1* together with a non-targeting control shRNA (Figures S1D and S1E). Knockdown of *WT1* by using a doxycycline-inducible shRNA (Figure 1D) led to a strong reduction of colony-forming ability in FLT3-ITD (Figure 1E) but not healthy $CD34^+$ cells (Figure 1F). *WT1* knockdown led to minor changes in the accessible chromatin landscape in the three samples, as measured by Assay for Transposase-Accessible Chromatin using Sequencing (ATAC-seq) (Figures 1G–1I), with the t(8;21) AML showing the greatest loss of ATAC-seq peaks (1,884 peaks) (Figure 1I). All primary AML cells displayed significant changes in gene expression that are described in further detail below.

Only a few model systems are available for studying t(8;21) AML, and the generation of patient-derived xenografts (PDXs) in mice from primary t(8;21) cells has proven to be very challenging (Sanchez et al., 2009; Sarry et al., 2011). We therefore used the Kasumi-1 cell line for elucidating the molecular details of how WT1 impacts leukemic growth, which is a well-characterized model of t(8;21) AML (Ptasinska et al., 2012, 2014, 2019). Similar to FLT3-ITD primary AML cells, and in concordance with the xenotransplantation experiments, Kasumi-1 cells displayed reduced cellular growth after transduction with two different *WT1* shRNA (Figures 1J and S1F) as well as reduced clonogenicity (Figure S1G) and slightly increased apoptosis (Figure S1H). However, we did not observe a difference in the cell cycle (Figure S1I). Taken together, our data demonstrate that WT1 is required for leukemic growth and clonogenicity in two different AML subtypes.

## Upregulation of the endogenous and mutated WT1 distorts growth

We next examined whether altering the expression levels of endogenous *WT1* would impact the growth of Kasumi-1 cells. To this end, we targeted a CRISPR-Cas9-histone acetyltransferase p300 (CRISPRa) fusion protein to potential regulatory regions of the endogenous *WT1* locus. *WT1* displays a number of DNase I hypersensitive sites (DHSs) that indicate potential *cis*-regulatory elements serving as guides to the design of guide RNAs. Each of these sites interacts with the promoter in leukemic cells, as determined by promoter capture Hi-C (CHi-C) analysis of primary cells (Figure 2A; Assi et al., 2019b). Previous work had identified an enhancer in intron 3 (Zhang et al., 1997) that we also detected as a DHS in AML cells and PBSCs. We also noticed DHSs in intron 8 and at the 3′ end of the gene that appeared to be specifically and reproducibly activated in primary AML but not in healthy cells (Figures 2A and S2A). One DHS in intron 8 was of interest because it contained motifs for the RUNX, ETS, and AP-1 family TFs (Figure 2B) that are important nodes in AML-subtype-specific GRNs. Targeting CRISPRa strongly increased histone H3K27 acetylation at DHSs (Figure 2C), but the DHS in intron 8 was the only site where targeting of CRISPRa led to an increase in WT1 protein and mRNA levels (Figures 2D and 2E), colony-forming ability (Figure 2F), and cellular growth (Figure 2G). Targeting did not lead to major changes in the cell cycle or apoptosis (Figures S2C and S2D). Based on its dynamic nature and the absence of the intron 8 DHS in normal $CD34^+$ cells, we speculate that this element responds to changes in the cellular environment associated with transformation, leading to an upregulation of *WT1* expression and increased growth, thus enforcing the leukemic phenotype. As a control, we also targeted CRISPRa to a CTCF-bound DHS downstream of WT1 to show that any changes in gene expression were not simply a consequence of forced histone acetylation at a site that interacts with the promoter.

WT1 is recurrently mutated in AML, and a number of different mutations in different exons have been detected (Figure 2H). We were interested to see whether these altered proteins would modulate the growth of t(8;21) cells as well. We therefore transduced Kasumi-1 cells with doxycycline-inducible constructs expressing proteins carrying frameshift mutations in exons 8 and 9 (Figure S2E) that truncated the protein at different $Zn^{2+}$-finger domains. The presence of the exon 8 mutation but not the exon 9 mutation led to significantly increased growth and clonogenicity and a decrease in apoptosis (Figure 2I; Figures S2G and S2H) but had no effect on the cell cycle (Figure S2F).

*WT1* is bound by RUNX1-ETO as well as RUNX1 at both the intron 3 and 8 sites (Figure S2B). It was previously shown that WT1 cooperates with RUNX1-ETO to cause a more aggressive AML in mice (Nishida et al., 2006). We therefore examined to what extent the expression of WT1 was dependent on the presence of RUNX1-ETO by knocking down its expression over a period of 10 days (Figure S2I). RUNX1-ETO depletion led to differentiation of such cells and a strong downregulation of *WT1* (Figure S2J). However, although RUNX1-ETO depletion on

(H) Schematic showing where frameshift mutations were introduced in WT1.
(I) Time course of cumulative cell growth of Kasumi-1 cells with or without doxycycline induction of WT1 mutant proteins carrying frameshift mutations in different exons.
For all panels, n = 3 biological replicates, and error bars show standard deviation. Significance was determined using a two-tailed Student's t test. *$p < 0.05$, **$p < 0.005$.

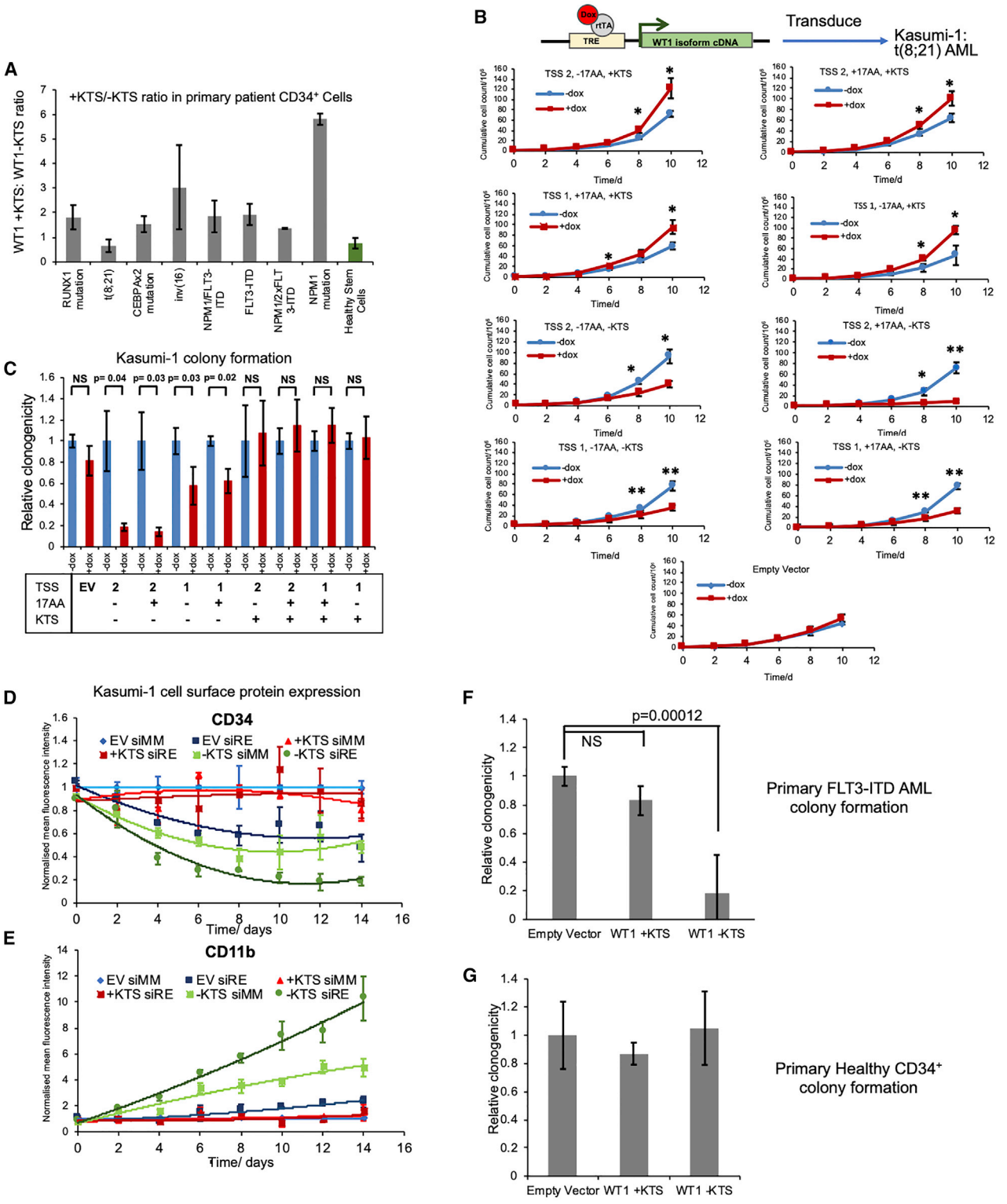



**Figure 3. Different WT1 isoforms have contrasting roles in AML maintenance**

(A) Ratio between expression of mRNAs encoding the *WT1* +KTS and *WT1* –KTS isoforms in purified primary patient CD34$^+$ cells.

(B and C) Top: Schematic showing the doxycycline-inducible strategy of *WT1* isoform expression. Kasumi-1 cells were transduced with doxycycline-inducible *WT1* isoforms. Bottom: Measurement of a time course of cumulative cell count (B) and of colony formation per 2,000 cells seeded in methylcellulose and normalized to empty vector control (C).

*(legend continued on next page)*

chromatin was complete in 2 days (Ptasinska et al., 2019), downregulation was delayed, indicating that *WT1* responds to changes in the GRN of differentiating cells rather than to the lack of RUNX1-ETO binding.

## WT1 contributes to oncogenesis in an isoform-specific fashion and cooperates with other oncogenes

Different AML subtypes within our patient cohort express different ratios of the *WT1* -KTS and *WT1* +KTS isoforms (Figure 3A). We therefore investigated whether the various WT1 isoforms and splice variants displayed a different ability to regulate leukemic growth in a leukemic context. We first transduced Kasumi-1 cells with constructs expressing various doxycycline-inducible WT1 isoforms (Figure S3A; Figure 3B, top panel), whereby we took great care to express the different proteins equally and at physiological levels. In growth assays, we did not see a difference between the long (promoter 1) and short (promoter 2) isoforms of WT1 but uncovered contrasting impacts of WT1 proteins with (+KTS) or without the splice variant lacking the KTS sequence (−KTS) (Figure 3B). All WT1 proteins carrying KTS stimulated growth, whereas all *WT1* –KTS isoforms suppressed growth. Although 17AA alternatively spliced isoforms have been described as having anti-apoptosis effects in a variety of cancer cell lines (Ito et al., 2006), we did not find such an effect in AML cells. The same was true for clonogenicity that strongly dropped after *WT1* –KTS isoform overexpression (Figure 3C). The *WT1* -KTS isoform mediated a slight increase in apoptosis (Figure S3B) and alterations in the cell cycle (Figure S3C). In accordance with these features, expression of the gene coding for the apoptosis regulator Caspase 9 *CASP9* was upregulated with both isoforms but was much stronger in *WT1* –KTS cells (Figure S3D).

Isoform expression differentially impacted on the differentiation stage of the cells, as measured by the analysis of the expression of the stem cell marker CD34 (*CD34*), the monocytic CD11b marker (*ITGAM*) (Figures 3D, 3E, and S3D), and of cellular morphology (Figure S3E). We also examined how the different isoforms cooperated with the transforming oncogene RUNX1-ETO by inducing the expression of WT1 proteins in the presence of a *RUNX1-ETO*-targeting small interfering RNA (siRNA) (siRE) as well as control siRNA (siMM) (Figures 3D and 3E). The depletion of RUNX1-ETO alone leads to the onset of cell differentiation indicated by a downregulation of CD34 and an upregulation of CD11b (Ptasinska et al., 2012). Again, induction of the two isoforms showed opposite effects; expression of the *WT1* –KTS isoform in RUNX1-ETO-depleted cells enhanced differentiation, whereas expression of the *WT1* +KTS isoform counteracted the effects of oncogene depletion.

We next tested these reagents in cultured primary AML cells from patients. The expression of the *WT1* –KTS isoform, but not the *WT1* +KTS isoform, had a strong impact on primary FLT3-ITD cells (Figures 3F and S3F), with a reduction in colony-forming ability, whereas normal cells were unaffected by the expression of either isoform (Figures 3G and S3H). We also used ATAC-seq to assess genome-wide chromatin accessibility after expressing different WT1 isoforms. The expression of the *WT1* –KTS isoform but not the *WT1* +KTS isoform led to altered patterns of chromatin accessibility in primary FLT3-ITD and t(8;21) AML (Figures S3I and S3J) but not in healthy CD34$^+$ cells (Figure S3K). In accordance with the previous result indicating increased myeloid differentiation, ATAC peaks gained in the presence of *WT1* –KTS show an increased enrichment of binding motifs for the myeloid master regulator PU.1 (Li et al., 2001).

In summary, our data demonstrate that the presence or absence of the KTS peptide and the ratio between the two isoforms have a profound impact not only on leukemic growth of AML cells but also on the chromatin landscape, cellular differentiation, and the cooperation with other oncoproteins.

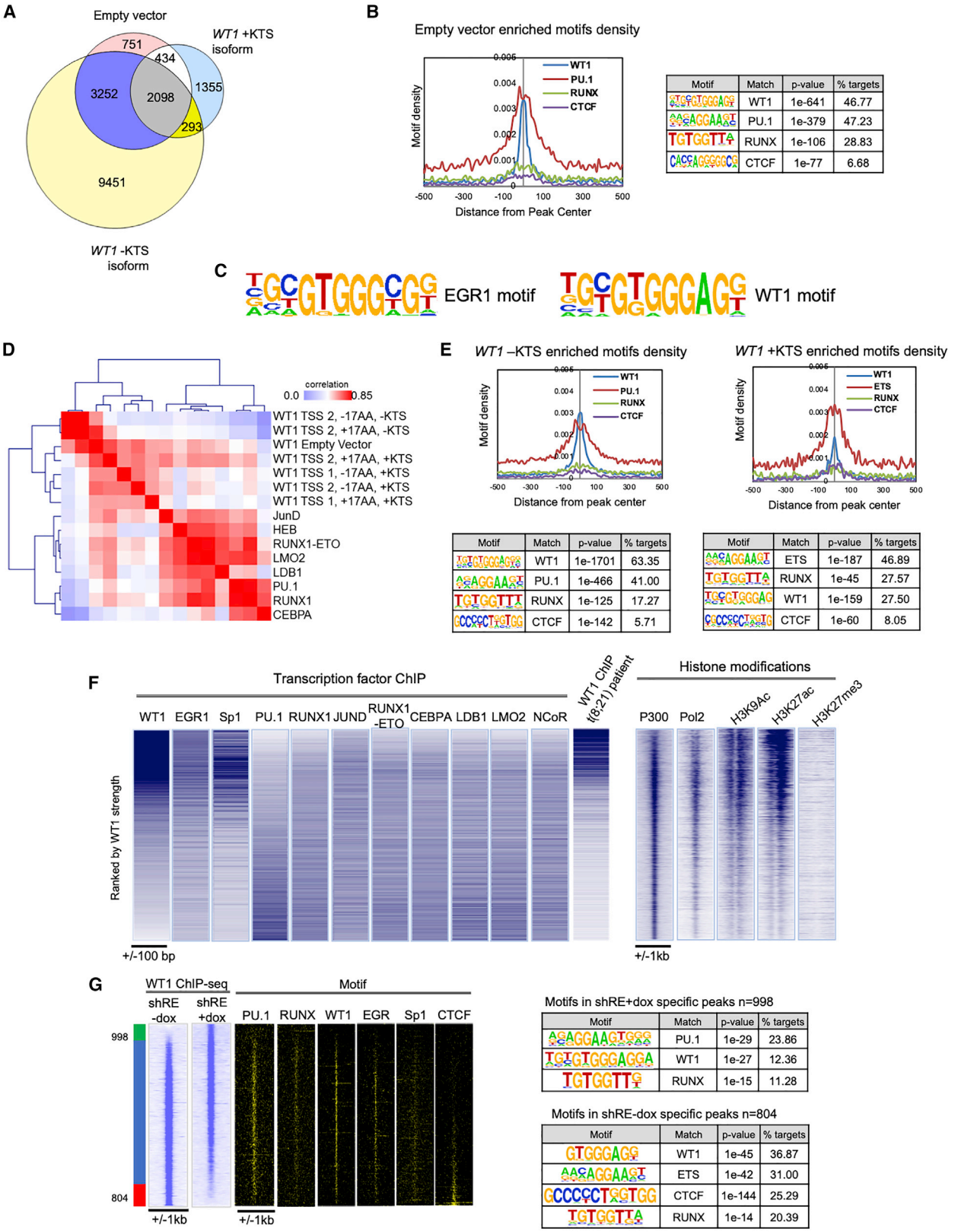
## The absence of the KTS peptide dramatically alters the global WT1 binding pattern

Our next experiments addressed the genomic locations with which WT1 interacted and whether the two isoforms bound to different genomic targets. To this end, we performed global chromatin immunoprecipitation sequencing (ChIP-seq) experiments with cells transduced with lentiviral vectors coding for *WT1* –KTS and *WT1* +KTS isoforms, in parallel with an empty vector control, to measure binding of the endogenous WT1 protein. The expression of the *WT1* –KTS isoform had no influence on the expression of the *WT1* +KTS isoform and vice versa (Data S1).

ChIP-seq revealed binding of WT1 isoforms to several thousand sites (Figure 4A; Figure S4A) with no isoform-specific preference for particular genomic regions. In concordance with its *in vitro* behavior (Laity et al., 2000), the *WT1* –KTS isoform interacted with many more genomic sites than the *WT1* +KTS isoform. This result was reproduced in cell clones carrying constructs expressing versions of the WT1 –KTS isoform with an hemagglutinin (HA)-tag fused their N terminus (Figure 4B; Figure S5A). We observed no difference in binding behavior between the long and short WT1 –KTS isoforms, and we were unable to perform ChIP on versions of WT1 carrying a C-terminal tag (data not shown). This result is in contrast to a previously published study that reported that WT1 +KTS bound to more binding sites than WT1 –KTS and that WT1 isoforms have a preference for different genomic locations (Ullmark et al., 2017). The differences are likely due to the better quality of our data.

WT1 peaks were enriched for ETS/PU.1 and RUNX1 motifs (Figure 4B) but sequences clustered apart from the sites bound by cell-fate-deciding factors, including RUNX1-ETO (Figure 4D; some data from Ptasinska et al., 2014, 2019). Based on these data, we could derive a chromatin-based position weight matrix

(D and E) Time course of mean fluorescence intensity changes of CD34 (D) and CD11b (E) surface marker expression in Kasumi-1 cells transduced with doxycycline-inducible empty vector, *WT1* +KTS, or *WT1* –KTS and transfected with siRNA against *RUNX1-ETO* (siRE) or a mismatch (siMM) control.
(F and G) Relative colony-formation ability of primary FLT3-ITD AML cells (F) or primary CD34$^+$ bone marrow cells from a healthy donor (G) transduced with doxycycline-inducible empty vector, *WT1* +KTS, or *WT1* –KTS.
For all panels, n = 3 biological replicates ,and error bars show standard deviation. Significance was determined using a two-tailed Student's t test. *$p < 0.05$, **$p < 0.005$.




*(legend on next page)*

(PWM) for WT1 in AML cells (Figure 4C). As suggested by previous *in vitro* experiments (Rauscher et al., 1990), this sequence is very similar to that of the EGR (early growth factor response) factor family with the exception of an obligatory adenine at position 9 that is more variable for EGR factors (but also includes the A as an infrequent base). The WT1 motif also shows some overlap with the Sp1 motif GGGG(C/A)GGGG. The enriched binding motif identified here was the same for the *WT1* –KTS and the *WT1* +KTS isoforms (Figure 4E).

To further characterize WT1-bound chromatin regions, we determined the binding profiles of EGR1 and Sp1 by ChIP-seq and aligned these signals with the WT1 peaks ranked according to WT1 binding signal, together with ChIP-seq peaks of other factors expressed in Kasumi-1 cells, for which data are available from our lab (Ptasinska et al., 2019) and others (Ben-Ami et al., 2013; Martens et al., 2012). On the same axis, we also aligned ChIP-seq signals measured in Kasumi-1 cells for aligned WT1 sites with RNA polymerase II (RNA Pol II); the co-activator p300; the Nuclear Receptor co-Repressor NCoR (Trombly et al., 2015); the histone modifications H3K27Ac, H3K9Ac (Ptasinska et al., 2019), and H3K27me3 (Trombly et al., 2015). We also determined the binding sites of the endogenous WT1 in primary cells of t(8;21) patients (Figure 4F) which demonstrated a strong overlap of WT1 binding sites in the Kasumi-1 cell line and in primary t(8;21) cells. WT1 binding sites were strongly associated with Sp1 binding sites but showed a weaker co-localization with EGR1 bound sites. WT1 co-localized with active (p300 associated, H3 acetylated) and with Pol-II-associated regions but not with polycomb (H3K27me3) and NCoR-bound chromatin. No preferred co-localization with any of the other measured factors was observed. We next examined the impact of the expression of the two WT1 isoforms on open chromatin patterns by using DNase I which gives a more refined cleavage pattern and sharper peaks than ATAC-seq (Karabacak Calviello et al., 2019). The expression of both isoforms changed the chromatin pattern and, interestingly, also the motif composition of enriched peaks (Figures S4D and S4E). In concordance with the activation of differentiation driven by the *WT1* –KTS isoform, we observed a strong enrichment of binding motifs for PU.1 in 1,742 open chromatin sites gained after induction (Figure S4D). This result was not observed with the *WT1* +KTS isoform, for which we observed a loss of PU.1 motifs in 3,131 accessible sites present before induction (Figure S4E). In addition, the onset of differentiation caused by the depletion of RUNX1-ETO after expression of an inducible shRNA caused a reduction of WT1 binding sites and a redistribution of WT1 sites toward *cis*-elements enriched in PU.1 motifs (Figure 4G; Figure S4F). Conversely, the overexpression of the *WT1* –KTS and *WT1* +KTS isoforms caused a shift of RUNX1-ETO binding toward sites with RUNX1 motifs (Figures S4F and S4G). Taken together, our analyses show that the binding pattern of WT1 is dependent on the presence of RUNX1-ETO and vice versa.

## WT1-isoform-specific regulation of AML-subtype-specific GRNs

We next determined the nature of gene expression changes in Kasumi-1 cells and in purified primary patient-derived blast cells after *WT1* knockdown and expression of *WT1* –KTS and *WT1* +KTS isoforms by using RNA sequencing (RNA-seq). Clustering of the RNA-seq profiles of differentially expressed genes showed extensive WT1-isoform- and cell-type-specific changes (Figure 5A; Figure S5A). *WT1* manipulation led to both up- and downregulation of gene expression, with approximately one-third of responsive genes being direct WT1 targets, which we assigned using our previously published CHi-C data (Ptasinska et al., 2019; Figure 5B). When we examined gene expression changes at WT1 target genes, the +KTS form showed more gene repression that the −KTS isoform, with its specific targets being mostly associated with gene activation (Figure S5A). Only a small number of direct WT1-bound genes were responding to WT1 knockdown (Figure 5C; Data S1). Interestingly, *WT1* knockdown and RUNX1-ETO knockdown changed gene expression in a similar fashion (Figure 5D), emphasizing that the two genes operate in the same pathway. KEGG pathway analysis (Kanehisa and Goto, 2000; Data S1) of the response to isoform overexpression demonstrated a strong response to overexpression of the *WT1* –KTS isoform, with multiple signaling pathways being upregulated, including differentiation-promoting cytokine receptors such as Granulocyte-Macrophage Stimulating Receptor (GM-CSFR$\alpha$ (*CSF2R*) or the main driver of macrophage differentiation *SPI1* (PU.1). Downregulated genes included *MYC*, which could explain the differentiation-promoting ability of this isoform. Overexpression of *WT1* +KTS affected fewer pathways, with the exception of a downregulation of genes involved in $Ca^{2+}$ transport and signaling (*CALM2*) encoding for Calmodulin or $Ca^{2+}$-stimulated phosphatases (PPTs). A number

**Figure 4. Identification and characterization of isoform-specific WT1-bound sites in chromatin**

(A) Venn diagram showing the overlap between WT1 ChIP-seq binding sites in Kasumi-1 cells transduced with doxycycline-inducible empty vector (endogenous WT1), *WT1* –KTS, or *WT1* +KTS.

(B) Average profile plots of motif density relative to distance from ChIP peak center in Kasumi-1 cells transduced with a doxycycline-inducible empty vector. A table of the frequencies of enriched TF binding motifs is shown alongside.

(C) *De-novo*-generated position weight matrix underlying WT1 binding together with that of EGR1 based on ChIP-seq experiments.

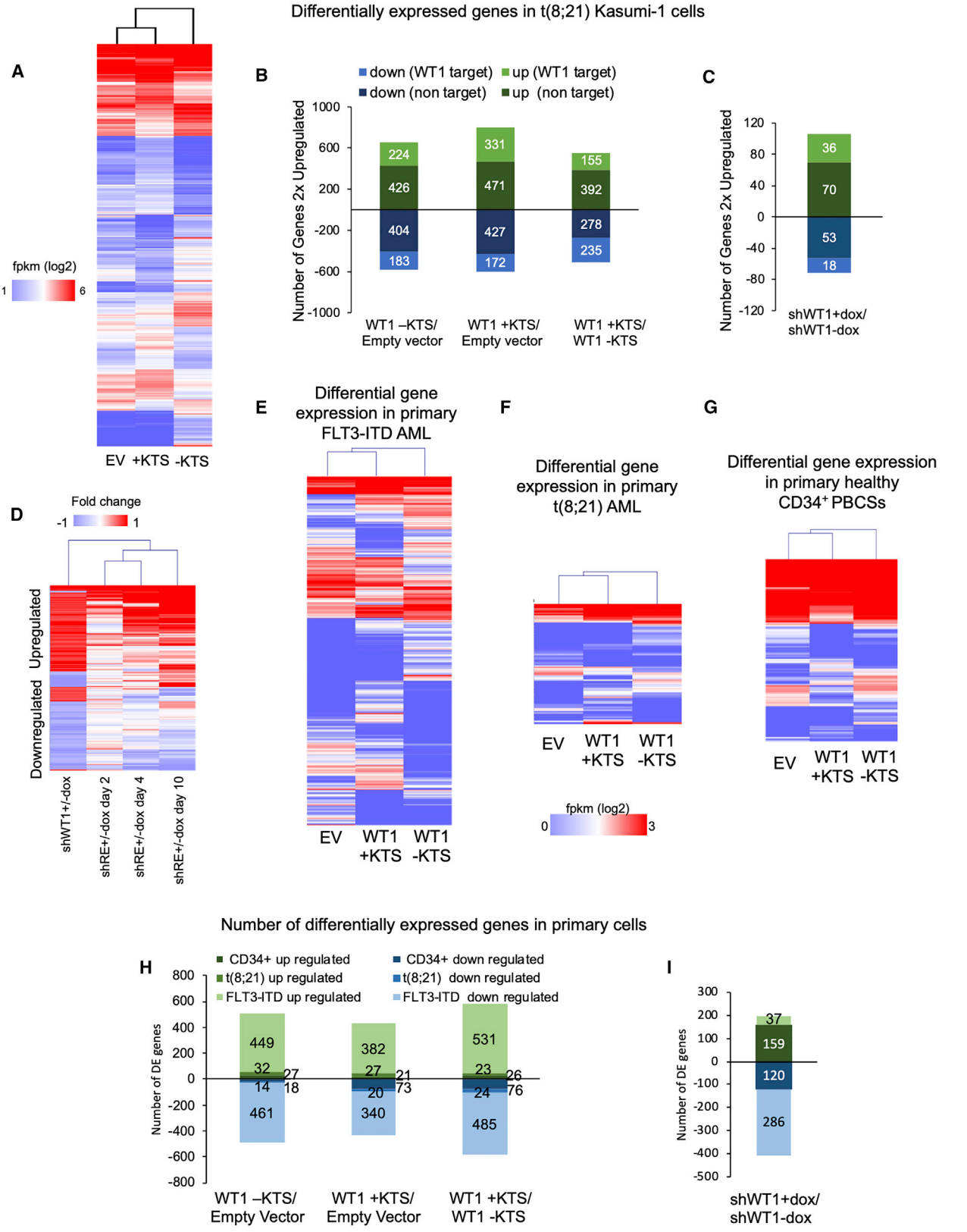
(D) Heatmap showing hierarchical clustering of Pearson correlation coefficients of WT1 isoform ChIP-seq sequence profiles and that of the indicated transcription factor ChIP-seq profiles.

(E) Average profile plots of motif density relative to distance from ChIP peak center in Kasumi-1 cells transduced with doxycycline-inducible *WT1* –KTS and *WT1* +KTS together with the table of enriched motifs in these sites.

(F) Density plots of the indicated TF and histone modification ChIP-seq peaks within a 200-bp window, ranked by tag count of WT1 peak size going from top to bottom, in Kasumi-1 cells. Also ranked beside them are WT1 ChIP-seq peaks in primary t(8;21) AML cells.

(G) Density plots of WT1 ChIP-seq peaks in a 2-kb window, ranked top to bottom by relative tag count of peaks in doxycycline-inducible shRNA against *RUNX1-ETO* relative to no-doxycycline control in Kasumi-1 cells. Plotted alongside are enriched motifs and tables of enriched motifs.

For all panels, significance was determined by a binomial test for motif enrichment analysis.




*(legend on next page)*

of responsive targets were shared by the two isoforms, but the effect of overexpression was not the same for all of them (Data S2).

The analyses of the responses of primary t(8;21) and FLT3-ITD AML cells to WT1 isoform overexpression identified both up- and downregulated genes, with multiple pathways being altered (Figures 5E–5G; Figure S5B; Data S3). Importantly, for both primary cells and transduced Kasumi-1 cells, each AML subtype responded to overexpression differently. Changes in gene expression were directly associated with differences in the accessible chromatin landscape (Figure S5C), with patterns clustering into the same groups. In primary cells, we found barely any overlap between cell types for the specific subsets of WT1-responsive genes, but the number of responding genes was higher in FLT3-ITD AML than in t(8;21) AML cells (Figures 5H and 5I; Data S3). However, the two AML subtypes shared some WT1-regulated pathways. The expression of genes involved in a number of signaling pathways was changed, including the interleukin-17 (IL-17) signaling pathway (Data S3), highlighting a reprogramming of the signaling environment towards inflammation (Iwakura et al., 2011).

The extensive collection of multi-omics data available from the Kasumi-1 cell line allowed us to construct a GRN of TF encoding genes, some of which are themselves subject to WT1 regulation (Figure S6, bottom half). This GRN can be used to link downstream non-TF target genes that are bound by WT1 and were more that 2-fold upregulated after *WT1* −KTS (Figure S6A) or *WT1* +KTS (Figure S6B) expression (Figure S6, top half). Some of these target genes encode potential drug targets, such as the PIM1 or FES kinases. These analyses clearly demonstrate that WT1-regulated genes are connected within a network of TF families that control signaling response (AP-1, ATF, IRF, FOX, STAT, and NFAT families), growth (MYC-family and HLH E-Box factors), hematopoiesis (RUNX1 and CBFβ), and myeloid differentiation (IRF, ETS, PU.1, and C/EBP TF families), highlighting the importance of extracellular signaling in understanding leukemic phenotypes and the formidable complexity of predicting the response to therapeutic intervention.

## WT1 binding is modulated by other TFs

Our data so far indicate that WT1 is part of a GRN that includes a number of signaling-responsive factors. The most interesting factor family in this context is the early growth-factor response (EGR) TF family that previously has been shown to respond to a number of different growth signals (Thiel and Cibelli, 2002) and whose binding site bears strong similarity to that of WT1 (Figure 4C). Bootstrapping analysis of binding motifs examining the significance of their co-localization within a distance of 50 bp demonstrated that WT1/EGR1 binding sites strongly overlapped with Sp1 binding sites (Figure 6A) which was confirmed by analyzing ChIP and motif density data (Figures 6B–6E). The analysis of the actual binding sequences unveiled the presence of a composite Sp/WT1/EGR motif with a defined spacing (Figure 6F) that could also be seen after a strand-specific analysis around the WT1/EGR1 peaks, suggesting that the factors act together by wrapping with their $Zn^{2+}$ fingers around opposite strands (Figure 6G; Figure S7A). This feature was not seen with other factors, such as ETS (Figure S7B). These data strongly suggest a complex interplay of the three factors at their respective genomic location, most of which are Pol-II-associated regions and promoters that will merit future investigation.

Given the similarity of binding sites (Figures 6A and 6C), we questioned whether EGR1 binding responded to changes in WT1 expression levels by performing EGR1 ChIP-seq assays in Kasumi-1 cells after *WT1* depletion (Figure 6H) and WT1 isoform overexpression (Figure 6I). Our experiment demonstrates that that WT1 perturbation yielded a complex response (Figures 6J and 6K). After WT1 knockdown, we observed a redistribution of EGR1 binding away from Sp1 sites as the cells differentiate, even at shared sites (Figure 6H). However, the effect of WT1 isoform overexpression was drastic; when ranked alongside peaks from cells transduced with an empty vector, more than 14,000 EGR1 peaks were lost after WT1 –KTS or +KTS isoform overexpression, with only 298 gained (Figure 6I), both at shared and non-shared sites, indicating that the two proteins regulate each other's binding patterns in different ways. This finding is exemplified by a *SOCS3* downstream DHS and the *CASP9* promoter (Figures S7C and S7D). In addition, consistent with its ability to drive cell differentiation, expression of *WT1* −KTS led to the appearance of new EGR1 sites associated with PU.1 motifs (Figure 6I).

Based on data from Assi et al. (2019b) that are derived from patients (digital DNase I footprinting and CHi-C data), we constructed an AML-specific sub-module of WT1 and EGR1-regulated genes in primary FLT3-ITD/NPM1 and primary t(8;21) AML (Figures S8A and S8B). This analysis showed that both *WT1* and *EGR1* are auto-regulated and that WT1- and EGR1-regulated genes are strongly interlinked with multiple targets containing motifs for both factors. This includes genes encoding important regulators, such as the apoptosis regulator BCL2 and the TF NFIX, which we showed to be an important network node for FLT3-ITD AML growth (Assi et al., 2019b). Another EGR1-bound gene is *FOXC1* that is an oncogene in its own right

**Figure 5. WT1 isoform expression produces distinct transcriptional signatures**

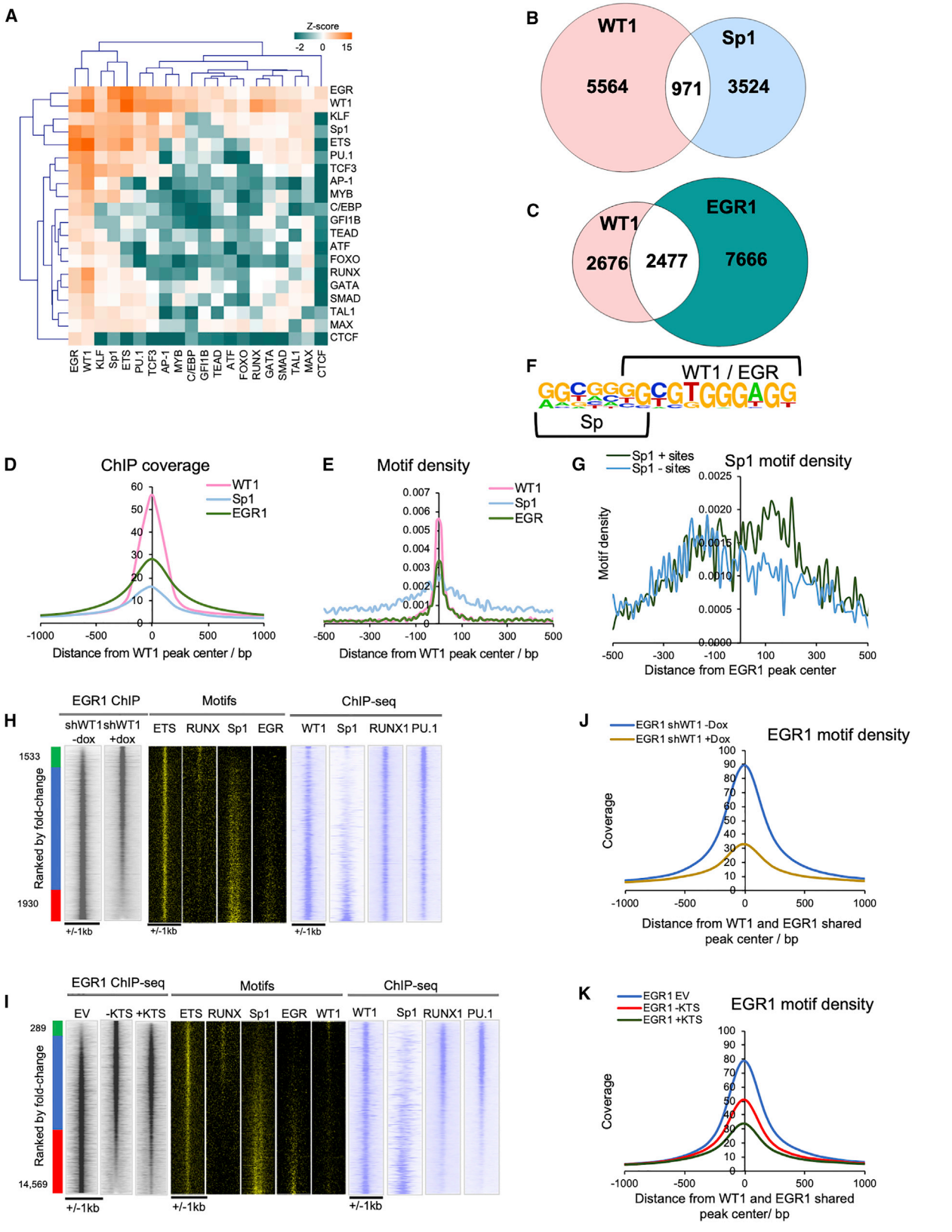
(A) Heatmap displaying hierarchical clustering of differentially expressed genes from the RNA-seq of Kasumi-1 cells transduced with a doxycycline-inducible empty vector, *WT1* +KTS, or *WT1* –KTS.

(B and C) Number of differentially expressed WT1 target or non-target genes from Kasumi-1 cells transduced with doxycycline-inducible empty vector, *WT1* +KTS, or *WT1* –KTS (B) or sh*WT1* (C).

(D) Heatmap with hierarchical clustering of differentially expressed genes (RNA-seq) from Kasumi-1 cells transduced with *RUNX1-ETO* measured over a time course of *RUNX1-ETO* depletion ranked alongside genes up- or downregulated after expression of doxycycline-inducible shRNA against *WT1*.

(E–G) Heatmap showing hierarchical clustering of differentially expressed genes in the RNA-seq profiles from primary FLT3-ITD-mutated AML (E), primary t(8;21) AML (F), and primary bone marrow $CD34^+$ cells from a healthy donor (G), transduced with doxycycline-inducible empty vector, *WT1* +KTS, or *WT1* –KTS constructs.

(H and I) Number of differentially expressed genes of the indicated cell types transduced with doxycycline-inducible empty vector, *WT1* +KTS, or *WT1* –KTS (H) or sh*WT1* constructs (I).




*(legend on next page)*

(Somerville et al., 2015) and that we also demonstrated to be an important FLT3-ITD AML network node (Assi et al., 2019b). The structure of this network suggested a signaling-dependent crosstalk between the two genes.

We next performed perturbation experiments to interrogate this module. In our primary patient FLT3-ITD AML culture, we conducted a number of perturbation experiments that show how oncogenic signaling and factor binding feed into WT1/EGR1 modules (Figures 7A–7C). We first treated primary cell cultures with the clinically approved FLT3 inhibitor Quizartinib (AC220) (Figure 7A) that repressed the expression of both genes. We also inhibited the RUNX1-CBFβ interaction with a specific small molecule (Figure 7B) that abrogated RUNX1 DNA binding (Illendula et al., 2016). Treatment led to a downregulation of *WT1* but not *EGR1*. Finally, we blocked all AP-1 activity by using a dominant-negative FOS peptide (dnFOS) (Olive et al., 1997) (Figure 7C) that again showed a strong downregulation of *WT1* but not *EGR1*. Taken together, our data show (1) the highly dynamic behavior of *WT1* and *EGR1* expression as well as chromatin binding in response to perturbation of each other's expression, (2) an important role of RUNX1 and AP-1 in regulating WT1 expression, and (3) an essential role of oncogenic signaling in *WT1*/*EGR1* upregulation.

## DISCUSSION

*WT1* overexpression is a marker for an aggressive type of AML independent of the AML subtype. It is upregulated after AML relapse and is a biomarker for relapse prediction, as it can be detected in minimal residual disease (MRD) (Inoue et al., 1994). In addition, *WT1* expression in MRD is a biomarker for AML relapse after post-allogeneic stem cell transplantation (Nomdedéu et al., 2018). Our work is therefore of clinical importance as it provides insights into how WT1 contributes to AML fitness and why it is upregulated. We find that (1) WT1 is an important component of AML maintenance within the GRNs of multiple subtypes of AML but not in normal cells; (2) it acts in an isoform-specific fashion; (3) specific mutants and isoforms, in particular the +KTS isoform, drive increased growth; (4) its overexpression causes aberrant growth, thus linking its expression to increased fitness; (5) the upregulation of its expression is linked to oncogenic signaling and early growth factor response gene action; and (6) most importantly, its binding within chromatin changes the expression of many different genes, but in an AML-subtype-specific fashion with little overlap in differentially expressed genes between AML subtypes.

Our ChIP and gene expression data allow us to propose models of AML-subtype-specific upregulation of *WT1* in response to external signals (Figures 7D and 7E). In t(8;21) cells, *WT1* is bound by RUNX1-ETO at the intron 3 enhancer and the potential negative regulatory element at intron 8. The intron 3 enhancer is bound by a number of TFs, including AP-1 family members, which, in turn, are dependent on the expression of the fusion protein and are required for the expression of multiple RUNX1-ETO targets (Frank et al., 1999; Martinez-Soria et al., 2019; Ptasinska et al., 2019). WT1 positively auto-regulates itself and is required for the elevated expression of *EGR1*, which is also a target of RUNX1-ETO, and is regulated by signaling in multiple contexts. EGR1 binds to *WT1* and thus is likely to modulate *WT1* expression in response to external signals as well, highlighting a complex interplay between the two factors. In t(8;21) cells, the knockdown of *RUNX1* expression has no effect on the expression of *WT1* (data from Ben-Ami et al., 2013). In contrast, *RUNX1* is part of the FLT3-ITD AML GRN, as its inhibition leads to *WT1* downregulation and its own expression is upregulated by the FLT3-ITD signal (Figure 7E), again demonstrating that the gene is differentially wired within different GRNs. In a similar vein, in FLT3-ITD AML, *WT1* and *EGR1* are bound by multiple signaling responsive factors such as AP-1 and STAT5 (Cauchy et al., 2015) and gene expression responds to oncogenic signaling. However, only *WT1* responds to perturbation by RUNX1/AP-1 inhibition in spite of *EGR1* being a target of both factors. Finally, *EGR1* strongly responds to WT1 knockdown, indicating that WT1 acts upstream of *EGR1*. We therefore conclude that elevated growth factor signaling in leukemic cells is a major reason for *WT1* upregulation and also provides an explanation for the worse prognosis in such patients, as WT1 itself drives the enhanced growth of leukemic cells. Indeed, activation of signaling plays a key role in growth and increasing fitness in AML. In apparently healthy individuals, the detection of mutations in signaling genes such as *KRAS*, *NRAS*, *NF1*, *JAK2*, and *CBL* could predict the development of AML around 6 years later (Abelson et al., 2018). Activating *KRAS* mutations have been demonstrated to confer increased competitive survival in xenotransplantation experiments (Burgess et al., 2017). Furthermore, the acquisition of a FLT3-ITD mutation in a NPM1 mutant background accelerates the relapse time as compared to the NPM1 mutation alone (Höllein et al., 2018). Finally, *WT1*

**Figure 6. The interplay of WT1, EGR1, and Sp1 zinc finger transcription factors**

(A) Bootstrapping analysis clustering the significance of co-localization of the indicated TF binding motifs within 50 bp within WT1 ChIP-seq peaks in Kasumi-1 cells.

(B) Venn diagram showing the overlap between WT1 and Sp1 ChIP-seq binding sites in Kasumi-1 cells.

(C) Venn diagram showing the overlap between WT1 and EGR1 ChIP-seq binding sites in Kasumi-1 cells

(D and E) Average profile plot of ChIP coverage (D) and motif density (E) relative to distance from the WT1 peak center in Kasumi-1 cells.

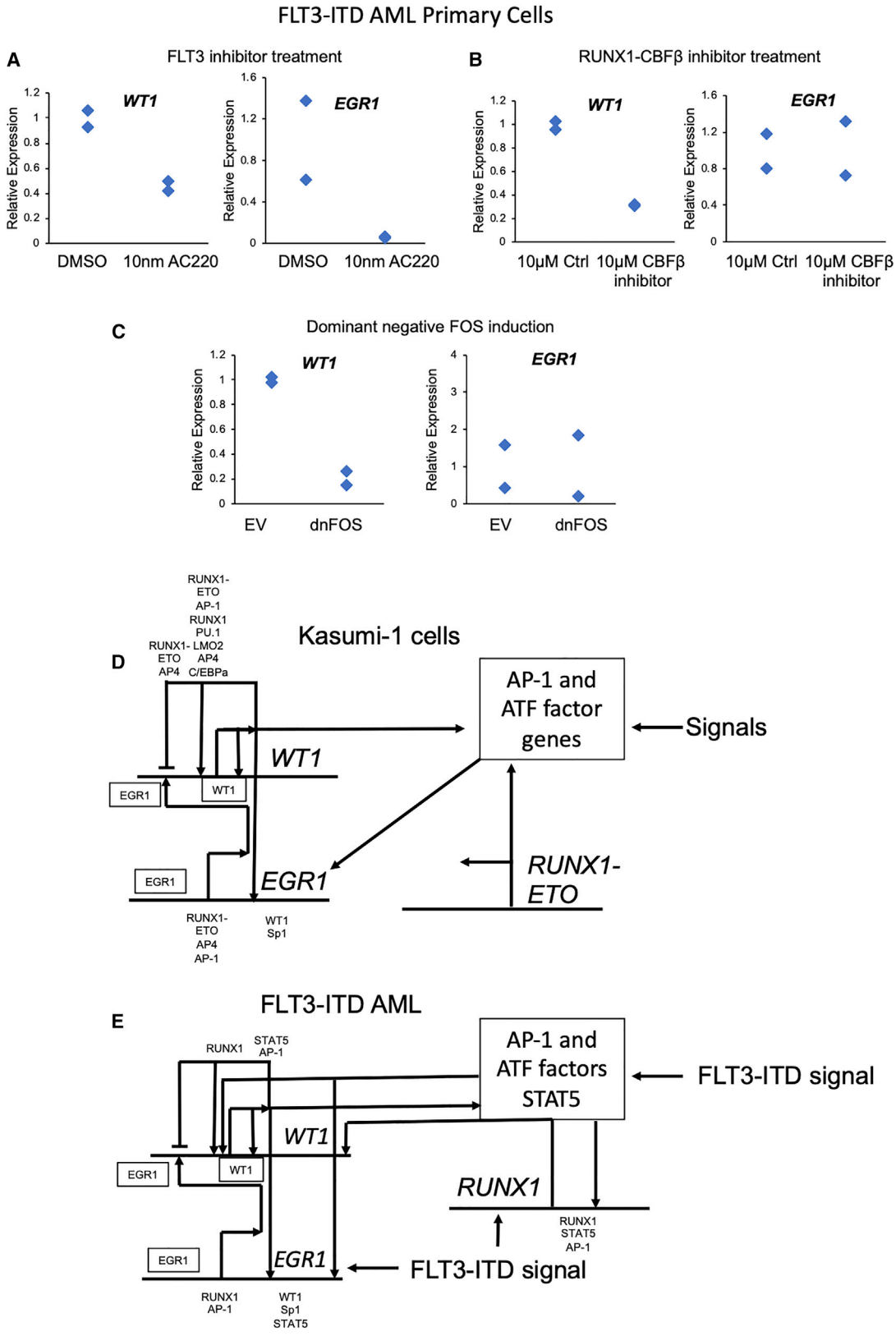
(F) Extended HOMER motif search in WT1 or EGR1 ChIP-seq experiments showing a composite Sp:WT1/EGR1 motif.

(G) Average profile plot of strand-specific Sp1 motif frequency relative to distance from the EGR1 peak summit, in EGR1 ChIP-seq peaks from Kasumi-1 cells (I).

(H and I) Density plots of EGR1 ChIP-seq peaks in a 2-kb window, ranked top to bottom by relative tag count of peaks in doxycycline-inducible shRNA against *WT1* relative to the no-doxycycline control in Kasumi-1 cells (H) or doxycycline-inducible *WT1* –KTS relative to the empty vector control (I). The bars on the side depict the number of specific and shared peaks. Plotted alongside are enriched TF binding motifs.

(J and K) Average profile plots of EGR1 ChIP coverage in Kasumi-1 cells transduced with sh*WT1* ± doxycycline (J) and Kasumi-1 cells transduced with empty vector, *WT1* –KTS, or *WT1* +KTS (K).

For all panels, significance was determined by a binomial test for motif enrichment analysis.




*(legend on next page)*

mutations have been described to cooperate with FLT3-ITD mutations in the development of a fully penetrant AML *in vivo* (Pronier et al., 2018), again alluding to the clonal advantages conferred by activated signaling.

Irrespective of the isoform, WT1 binding sites overlap with RNA Pol-II-bound regions, indicating that it is intricately associated with mRNA synthesis. WT1 depletion leads to AML-subtype-specific changes in the *cis*-regulatory landscape. The molecular basis for these changes is the regulatory interplay of WT1 with other TFs, in particular EGR family members. In t(8;21) cells, WT1 binding responds to the depletion of RUNX1-ETO, and the knockdown of either protein leads to similar changes in the gene expression pattern, suggesting that both reinforce each other's action. WT1 can therefore be seen as an essential part of an AML maintenance network. Our analysis shows the nature and structure of such networks and how they are connected. However, as outlined in the introduction, WT1 is likely to regulate gene expression at multiple levels, including mRNA processing and stability. In this context, it is interesting to note that WT1 binds to its own intron-exon boundaries (Figure S2B), including the 3′ UTR. Experiments outside the scope of this study are currently examining whether WT1 impacts splicing and mRNA turnover of its own and other genes.

Although TFs were long considered "undruggable," recent studies have pointed toward several potential therapeutic strategies targeting their activities. Direct inhibition of TF complexes is exemplified by CBFβ-RUNX1 complex inhibition (Illendula et al., 2016; Figure 7B). An alternate strategy is to target encoding upstream regulators (Figures 7D and 7E) or downstream effector genes (Figure S6, top half). Our findings demonstrate that WT1 would present an important subtype-independent therapeutic target. The success of using T cell receptor therapies against WT1, particularly in the post-allogeneic stem cell transplant setting (Chapuis et al., 2019), raises expectations that its clinical use would be feasible. However, our finding that different WT1 isoforms have contrasting biological activities in AML indicate that caution is needed when designing therapeutic approaches that may shift the isoform ratio. Our results also highlight the need to account for the AML subtype, as the genomic response to WT1 overexpression and *WT1* regulation differs between subtypes, meriting efforts to understand how their signaling network is rewired to keep AML cells alive and how this feeds into the WT1/EGR1 axis. The success of choosing the right target will depend on a thorough understanding on the role of aberrantly expressed/mutated TFs within a transformed GRN.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell lines and cell line culture
  - Patient samples, purification and culture
- METHOD DETAILS
  - Targeted myeloid sequencing panel
  - RNA Sequencing
  - DNase-I hypersensitive site mapping
  - Assay for transposase accessible chromatin (ATAC)-seq
  - Chromatin immunoprecipitation (ChIP)-seq
  - Transcription factor and gene regulatory network identification
  - siRNA knockdown
  - Cloning of shRNA constructs
  - Cloning of overexpression constructs
  - WT1 mutant cloning
  - HA tagged WT1 cloning
  - CRISPR activation
  - Lentiviral transduction
  - Colony formation assays
  - Growth curves of cell lines
  - Flow cytometry
  - Cytospins
  - Western blot
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.109010.

## ACKNOWLEDGMENTS

We are grateful to the Genomics Birmingham Facility and to the University of Birmingham Flow Cytometry platform for their expert services. Furthermore, we thank Prof. Nick Hastie and Dr. Ruthrothaselvi Bharathavikru at the

**Figure 7. *WT1* and *EGR1* regulatory modules in t(8;21) and FLT3/NPM1 AML are connected to signaling processes**

(A) Dot plot of relative gene expression as determined by RNA-seq of WT1 and EGR1 of primary FLT3-ITD cells with treatment with DMSO or the FLT3-ITD inhibitor Quizartinib (AC220). Note that all experiments were done twice.

(B) Dot plot of relative gene expression as determined by RNA-seq of *WT1* and *EGR1* in cultured primary FLT3-ITD cells after treatment with a control compound or the RUNX1-CBFβ inhibitor (Illendula et al., 2016).

(C) Dot plot of relative gene expression of *WT1* and *EGR1* of primary FLT3-ITD cells transduced with doxycycline-inducible empty vector or dominant-negative FOS (dnFOS), as measured by RT-PCR.

(D) *WT1* and *EGR1* regulatory subnetwork specific for Kasumi-1 cells based on the ChIP data described here and elsewhere (Ben-Ami et al., 2013; Martens et al., 2012; Ptasinska et al., 2014), as well as perturbation experiments (dnFOS expression, *RUNX1-ETO* knockdown, *WT1* knockdown, and *RUNX1* knockdown; Ben-Ami et al., 2013). *RUNX1-ETO* expression upregulates AP-1 (Martinez-Soria et al., 2019); *RUNX1* knockdown has no effect on WT1 expression, and WT1 knockdown slightly reduces EGR1 expression (this study).

(E) *WT1* and *EGR1* regulatory subnetwork specific for *FLT3-ITD/NPM1*-mutated AML based partly on data from Assi et al. (2019b), STAT5 ChIP data from Cauchy et al. (2015), and the perturbation and ChIP experiments described here.

University of Edinburgh for advice on WT1 ChIP. This work was funded by a Medical Research Council and Leukaemia UK Clinical Research Training Fellowship to S.P. and by grants from Blood Cancer UK (specialist program grant 15001) and from the Medical Research Council (MR/S021469/1) to C.B. and P.N.C.

## AUTHOR CONTRIBUTIONS

S.P. designed the study, performed experiments, and wrote the paper. S.A.A. performed bioinformatics analyses. P.S.C., D.J.L.C., and A.P. performed experiments. S.M. and N.S. provided reagents. O.H. and P.N.C. supervised experiments. C.B. conceived, designed, and supervised the study and wrote the paper.

## DECLARATION OF INTERESTS

The authors declare no competing financial interests.

Received: August 7, 2020
Revised: February 4, 2021
Accepted: March 26, 2021
Published: April 20, 2021

## REFERENCES

Abelson, S., Collord, G., Ng, S.W.K., Weissbrod, O., Mendelson Cohen, N., Niemeyer, E., Barda, N., Zuzarte, P.C., Heisler, L., Sundaravadanam, Y., et al. (2018). Prediction of acute myeloid leukaemia risk in healthy individuals. Nature *559*, 400–404.

Alanazi, B., Munje, C.R., Rastogi, N., Williamson, A.J.K., Taylor, S., Hole, P.S., Hodges, M., Doyle, M., Baker, S., Gilkes, A.F., et al. (2020). Integrated nuclear proteomics and transcriptomics identifies S100A4 as a therapeutic target in acute myeloid leukemia. Leukemia *34*, 427–440.

Assi, S.A., Bonifer, C., and Cockerill, P.N. (2019a). Rewiring of the Transcription Factor Network in Acute Myeloid Leukemia. Cancer Inform. *18*, 1176935119859863.

Assi, S.A., Imperato, M.R., Coleman, D.J.L., Pickin, A., Potluri, S., Ptasinska, A., Chin, P.S., Blair, H., Cauchy, P., James, S.R., et al. (2019b). Subtype-specific regulatory network rewiring in acute myeloid leukemia. Nat. Genet. *51*, 151–162.

Barragán, E., Cervera, J., Bolufer, P., Ballester, S., Martín, G., Fernández, P., Collado, R., Sayas, M.J., and Sanz, M.A. (2004). Prognostic implications of Wilms' tumor gene (WT1) expression in patients with de novo acute myeloid leukemia. Haematologica *89*, 926–933.

Becker, H., Marcucci, G., Maharry, K., Radmacher, M.D., Mrózek, K., Margeson, D., Whitman, S.P., Paschka, P., Holland, K.B., Schwind, S., et al. (2010). Mutations of the Wilms tumor 1 gene (WT1) in older patients with primary cytogenetically normal acute myeloid leukemia: a Cancer and Leukemia Group B study. Blood *116*, 788–792.

Ben-Ami, O., Friedman, D., Leshkowitz, D., Goldenberg, D., Orlovsky, K., Pencovich, N., Lotem, J., Tanay, A., and Groner, Y. (2013). Addiction of t(8;21) and inv(16) acute myeloid leukemia to native RUNX1. Cell Rep. *4*, 1131–1143.

Bert, A.G., Johnson, B.V., Baxter, E.W., and Cockerill, P.N. (2007). A modular enhancer is differentially regulated by GATA and NFAT elements that direct different tissue-specific patterns of nucleosome positioning and inducible chromatin remodeling. Mol. Cell Biol. *27*, 2870–2885.

Bickmore, W.A., Oghene, K., Little, M.H., Seawright, A., van Heyningen, V., and Hastie, N.D. (1992). Modulation of DNA binding specificity by alternative splicing of the Wilms tumor wt1 gene transcript. Science *257*, 235–237.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.H., Pagès, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics *25*, 1091–1093.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Burgess, M.R., Hwang, E., Mroue, R., Bielski, C.M., Wandler, A.M., Huang, B.J., Firestone, A.J., Young, A., Lacap, J.A., Crocker, L., et al. (2017). KRAS Allelic Imbalance Enhances Fitness and Modulates MAP Kinase Dependence in Cancer. Cell *168*, 817–829.e15.

Call, K.M., Glaser, T., Ito, C.Y., Buckler, A.J., Pelletier, J., Haber, D.A., Rose, E.A., Kral, A., Yeger, H., Lewis, W.H., et al. (1990). Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. Cell *60*, 509–520.

Cauchy, P., James, S.R., Zacarias-Cabeza, J., Ptasinska, A., Imperato, M.R., Assi, S.A., Piper, J., Canestraro, M., Hoogenkamp, M., Raghavan, M., et al. (2015). Chronic FLT3-ITD Signaling in Acute Myeloid Leukemia Is Connected to a Specific Chromatin Signature. Cell Rep. *12*, 821–836.

Chapuis, A.G., Egan, D.N., Bar, M., Schmitt, T.M., McAfee, M.S., Paulson, K.G., Voillet, V., Gottardo, R., Ragnarsson, G.B., Bleakley, M., et al. (2019). T cell receptor gene therapy targeting WT1 prevents acute myeloid leukemia relapse post-transplant. Nat. Med. *25*, 1064–1072.

Chau, Y.-Y., Brownstein, D., Mjoseng, H., Lee, W.-C., Buza-Vidas, N., Nerlov, C., Jacobsen, S.E., Perry, P., Berry, R., Thornburn, A., et al. (2011). Acute multiple organ failure in adult mice deleted for the developmental regulator Wt1. PLoS Genet. *7*, e1002404.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat. Methods *14*, 959–962.

Davies, R.C., Calvio, C., Bratt, E., Larsson, S.H., Lamond, A.I., and Hastie, N.D. (1998). WT1 interacts with the splicing factor U2AF65 in an isoform-dependent manner and can be incorporated into spliceosomes. Genes Dev. *12*, 3217–3225.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dow, L.E., Premsrirut, P.K., Zuber, J., Fellmann, C., McJunkin, K., Miething, C., Park, Y., Dickins, R.A., Hannon, G.J., and Lowe, S.W. (2012). A pipeline for the generation of shRNA transgenic mice. Nat. Protoc. *7*, 374–393.

Erickson, P., Gao, J., Chang, K.S., Look, T., Whisenant, E., Raimondi, S., Lasher, R., Trujillo, J., Rowley, J., and Drabkin, H. (1992). Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. Blood *80*, 1825–1831.

Frank, R.C., Sun, X., Berguido, F.J., Jakubowiak, A., and Nimer, S.D. (1999). The t(8;21) fusion protein, AML1/ETO, transforms NIH3T3 cells and activates AP-1. Oncogene *18*, 1701–1710.

Goode, D.K., Obier, N., Vijayabaskar, M.S., Lie-A-Ling, M., Lilly, A.J., Hannah, R., Lichtinger, M., Batta, K., Florkowska, M., Patel, R., et al. (2016). Dynamic Gene Regulatory Networks Drive Hematopoietic Specification and Differentiation. Dev. Cell *36*, 572–587.

Hamilton, T.B., Barilla, K.C., and Romaniuk, P.J. (1995). High affinity binding sites for the Wilms' tumour suppressor protein WT1. Nucleic Acids Res. *23*, 277–284.

Hayakawa, F., Towatari, M., Kiyoi, H., Tanimoto, M., Kitamura, T., Saito, H., and Naoe, T. (2000). Tandem-duplicated Flt3 constitutively activates STAT5 and MAP kinase and introduces autonomous cell growth in IL-3-dependent cell lines. Oncogene *19*, 624–631.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Höllein, A., Meggendorfer, M., Dicker, F., Jeromin, S., Nadarajah, N., Kern, W., Haferlach, C., and Haferlach, T. (2018). *NPM1* mutated AML can relapse with wild-type *NPM1*: persistent clonal hematopoiesis can drive relapse. Blood Adv. *2*, 3118–3125.

Hou, H.-A., Huang, T.-C., Lin, L.-I., Liu, C.-Y., Chen, C.-Y., Chou, W.-C., Tang, J.-L., Tseng, M.-H., Huang, C.-F., Chiang, Y.-C., et al. (2010). WT1 mutation in

470 adult patients with acute myeloid leukemia: stability during disease evolution and implication of its incorporation into a survival scoring system. Blood *115*, 5222–5231.

Illendula, A., Gilmour, J., Grembecka, J., Tirumala, V.S.S., Boulton, A., Kuntimaddi, A., Schmidt, C., Wang, L., Pulikkan, J.A., Zong, H., et al. (2016). Small Molecule Inhibitor of CBFβ-RUNX Binding for RUNX Transcription Factor Driven Cancers. EBioMedicine *8*, 117–131.

Inoue, K., Sugiyama, H., Ogawa, H., Nakagawa, M., Yamagami, T., Miwa, H., Kita, K., Hiraoka, A., Masaoka, T., Nasu, K., et al. (1994). WT1 as a new prognostic factor and a new marker for the detection of minimal residual disease in acute leukemia. Blood *84*, 3071–3079.

Ito, K., Oji, Y., Tatsumi, N., Shimizu, S., Kanai, Y., Nakazawa, T., Asada, M., Jomgeow, T., Aoyagi, S., Nakano, Y., et al. (2006). Antiapoptotic function of 17AA(+)WT1 (Wilms' tumor gene) isoforms on the intrinsic apoptosis pathway. Oncogene *25*, 4217–4229.

Iwakura, Y., Ishigame, H., Saijo, S., and Nakae, S. (2011). Functional specialization of interleukin-17 family members. Immunity *34*, 149–162.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30.

Karabacak Calviello, A., Hirsekorn, A., Wurmus, R., Yusuf, D., and Ohler, U. (2019). Reproducible inference of transcription factor footprints in ATAC-seq and DNase-seq datasets using protocol-specific bias modeling. Genome Biol. *20*, 42.

Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., and Kent, W.J. (2004). The UCSC Table Browser data retrieval tool. Nucleic Acids Res. *32*, D493–D496.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res. *12*, 996–1006.

King-Underwood, L., Renshaw, J., and Pritchard-Jones, K. (1996). Mutations in the Wilms' tumor gene WT1 in leukemias. Blood *87*, 2171–2179.

Kreidberg, J.A., Sariola, H., Loring, J.M., Maeda, M., Pelletier, J., Housman, D., and Jaenisch, R. (1993). WT-1 is required for early kidney development. Cell *74*, 679–691.

Laity, J.H., Chung, J., Dyson, H.J., and Wright, P.E. (2000). Alternative splicing of Wilms' tumor suppressor protein modulates DNA binding activity through isoform-specific DNA-induced conformational changes. Biochemistry *39*, 5341–5348.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Larsson, S.H., Charlieu, J.P., Miyagawa, K., Engelkamp, D., Rassoulzadegan, M., Ross, A., Cuzin, F., van Heyningen, V., and Hastie, N.D. (1995). Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing. Cell *81*, 391–401.

Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A., Hoadley, K., Triche, T.J., Jr., Laird, P.W., Baty, J.D., et al.; Cancer Genome Atlas Research Network (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. *368*, 2059–2074.

Li, Y., Okuno, Y., Zhang, P., Radomska, H.S., Chen, H., Iwasaki, H., Akashi, K., Klemsz, M.J., McKercher, S.R., Maki, R.A., and Tenen, D.G. (2001). Regulation of the PU.1 gene by distal elements. Blood *98*, 2958–2965.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Martens, J.H.A., Mandoli, A., Simmer, F., Wierenga, B.-J., Saeed, S., Singh, A.A., Altucci, L., Vellenga, E., and Stunnenberg, H.G. (2012). ERG and FLI1 binding sites demarcate targets for aberrant epigenetic regulation by AML1-ETO in acute myeloid leukemia. Blood *120*, 4038–4048.

Martinez-Soria, N., McKenzie, L., Draper, J., Ptasinska, A., Issa, H., Potluri, S., Blair, H.J., Pickin, A., Isa, A., Chin, P.S., et al. (2019). The Oncogenic Transcription Factor RUNX1/ETO Corrupts Cell Cycle Regulation to Drive Leukemic Transformation. Cancer Cell *35*, 705.

Mifsud, B., Martincorena, I., Darbo, E., Sugar, R., Schoenfelder, S., Fraser, P., and Luscombe, N.M. (2017). GOTHiC, a probabilistic model to resolve complex biases and to identify real interactions in Hi-C data. PLoS One *12*, e0174744.

Moriya, S., Takiguchi, M., and Seki, N. (2008). Expression of the WT1 gene -KTS domain isoforms suppresses the invasive ability of human lung squamous cell carcinoma cells. Int. J. Oncol. *32*, 349–356.

Mostoslavsky, G., Kotton, D.N., Fabian, A.J., Gray, J.T., Lee, J.-S., and Mulligan, R.C. (2005). Efficiency of transduction of highly purified murine hematopoietic stem cells by lentiviral and oncoretroviral vectors under conditions of minimal in vitro manipulation. Mol. Ther. *11*, 932–940.

Nishida, S., Hosen, N., Shirakata, T., Kanato, K., Yanagihara, M., Nakatsuka, S., Hoshida, Y., Nakazawa, T., Harada, Y., Tatsumi, N., et al. (2006). AML1-ETO rapidly induces acute myeloblastic leukemia in cooperation with the Wilms tumor gene, WT1. Blood *107*, 3303–3312.

Nomdedéu, J.F., Esquirol, A., Carricondo, M., Pratcorona, M., Hoyos, M., Garrido, A., Rubio, M., Bussaglia, E., García-Cadenas, I., Estivill, C., et al. (2018). Bone Marrow WT1 Levels in Allogeneic Hematopoietic Stem Cell Transplantation for Acute Myelogenous Leukemia and Myelodysplasia: Clinically Relevant Time Points and 100 Copies Threshold Value. Biol. Blood Marrow Transplant. *24*, 55–63.

Olive, M., Krylov, D., Echlin, D.R., Gardner, K., Taparowsky, E., and Vinson, C. (1997). A dominant negative to activation protein-1 (AP1) that abolishes DNA binding and inhibits oncogenesis. J. Biol. Chem. *272*, 18586–18594.

Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V.I., Paschka, P., Roberts, N.D., Potter, N.E., Heuser, M., Thol, F., Bolli, N., et al. (2016). Genomic Classification and Prognosis in Acute Myeloid Leukemia. N. Engl. J. Med. *374*, 2209–2221.

Pronier, E., Bowman, R.L., Ahn, J., Glass, J., Kandoth, C., Merlinsky, T.R., Whitfield, J.T., Durham, B.H., Gruet, A., Hanasoge Somasundara, A.V., et al. (2018). Genetic and epigenetic evolution as a contributor to WT1-mutant leukemogenesis. Blood *132*, 1265–1278.

Ptasinska, A., Assi, S.A., Mannari, D., James, S.R., Williamson, D., Dunne, J., Hoogenkamp, M., Wu, M., Care, M., McNeill, H., et al. (2012). Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding. Leukemia *26*, 1829–1841.

Ptasinska, A., Assi, S.A., Martinez-Soria, N., Imperato, M.R., Piper, J., Cauchy, P., Pickin, A., James, S.R., Hoogenkamp, M., Williamson, D., et al. (2014). Identification of a dynamic core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal. Cell Rep. *8*, 1974–1988.

Ptasinska, A., Pickin, A., Assi, S.A., Chin, P.S., Ames, L., Avellino, R., Gröschel, S., Delwel, R., Cockerill, P.N., Osborne, C.S., and Bonifer, C. (2019). RUNX1-ETO Depletion in t(8;21) AML Leads to C/EBPα- and AP-1-Mediated Alterations in Enhancer-Promoter Interaction. Cell Rep. *28*, 3022–3031.e7.

Quinlan, A.R. (2014). BEDTools: The Swiss-Army Tool for Genome Feature Analysis. Curr. Protoc. Bioinformatics *47*, 11.12.11-34.

Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. *42*, W187–W191.

Rampal, R., Alkalin, A., Madzo, J., Vasanthakumar, A., Pronier, E., Patel, J., Li, Y., Ahn, J., Abdel-Wahab, O., Shih, A., et al. (2014). DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. Cell Rep. *9*, 1841–1855.

Rauscher, F.J., III, Morris, J.F., Tournay, O.E., Cook, D.M., and Curran, T. (1990). Binding of the Wilms' tumor locus zinc finger protein to the EGR-1 consensus sequence. Science *250*, 1259–1262.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

Saldanha, A.J. (2004). Java Treeview—extensible visualization of microarray data. Bioinformatics *20*, 3246–3248.

Sanchez, P.V., Perry, R.L., Sarry, J.E., Perl, A.E., Murphy, K., Swider, C.R., Bagg, A., Choi, J.K., Biegel, J.A., Danet-Desnoyers, G., and Carroll, M. (2009). A robust xenotransplantation model for acute myeloid leukemia. Leukemia *23*, 2109–2117.

Sarry, J.-E., Murphy, K., Perry, R., Sanchez, P.V., Secreto, A., Keefer, C., Swider, C.R., Strzelecki, A.-C., Cavelier, C., Récher, C., et al. (2011). Human acute myelogenous leukemia stem cells are rare and heterogeneous when assayed in NOD/SCID/IL2Rγc-deficient mice. J. Clin. Invest. *121*, 384–395.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

Somerville, T.D., Wiseman, D.H., Spencer, G.J., Huang, X., Lynch, J.T., Leong, H.S., Williams, E.L., Cheesman, E., and Somervaille, T.C. (2015). Frequent Derepression of the Mesenchymal Transcription Factor Gene FOXC1 in Acute Myeloid Leukemia. Cancer Cell *28*, 329–342.

Tate, J.G., Bamford, S., Jubb, H.C., Sondka, Z., Beare, D.M., Bindal, N., Boutselakis, H., Cole, C.G., Creatore, C., Dawson, E., et al. (2019). COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res. *47*, D941–D947.

Thiel, G., and Cibelli, G. (2002). Regulation of life and death by the zinc finger transcription factor Egr-1. J. Cell. Physiol. *193*, 287–292.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat. Biotechnol. *31*, 46–53.

Trincado, J.L., Entizne, J.C., Hysenaj, G., Singh, B., Skalic, M., Elliott, D.J., and Eyras, E. (2018). SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. Genome Biol. *19*, 40.

Trombly, D.J., Whitfield, T.W., Padmanabhan, S., Gordon, J.A., Lian, J.B., van Wijnen, A.J., Zaidi, S.K., Stein, J.L., and Stein, G.S. (2015). Genome-wide co-occupancy of AML1-ETO and N-CoR defines the t(8;21) AML signature in leukemic cells. BMC Genomics *16*, 309.

Tuna, M., Chavez-Reyes, A., and Tari, A.M. (2005). HER2/neu increases the expression of Wilms' Tumor 1 (WT1) protein to stimulate S-phase proliferation and inhibit apoptosis in breast cancer cells. Oncogene *24*, 1648–1652.

Ullmark, T., Järvstråt, L., Sandén, C., Montano, G., Jernmark-Nilsson, H., Lilljebjörn, H., Lennartsson, A., Fioretos, T., Drott, K., Vidovic, K., et al. (2017). Distinct global binding patterns of the Wilms tumor gene 1 (WT1) -KTS and +KTS isoforms in leukemic cells. Haematologica *102*, 336–345.

Wang, Y., Xiao, M., Chen, X., Chen, L., Xu, Y., Lv, L., Wang, P., Yang, H., Ma, S., Lin, H., et al. (2015). WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. Mol. Cell *57*, 662–673.

Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F., Xia, J., et al. (2012). The origin and evolution of mutations in acute myeloid leukemia. Cell *150*, 264–278.

Wingett, S., Ewels, P., Furlan-Magaril, M., Nagano, T., Schoenfelder, S., Fraser, P., and Andrews, S. (2015). HiCUP: pipeline for mapping and processing Hi-C data. F1000Res. *4*, 1310.

Zhang, X., Xing, G., Fraizer, G.C., and Saunders, G.F. (1997). Transactivation of an intronic hematopoietic-specific enhancer of the human Wilms' tumor 1 gene by GATA-1 and c-Myb. J. Biol. Chem. *272*, 29272–29280.

Zhang, H.-M., Chen, H., Liu, W., Liu, H., Gong, J., and Wang, H. (2011). AnimalTFDB: a comprehensive animal transcription factor database. Nucleic Acids Res. https://doi.org/10.1093/nar/gkr965.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| Mouse anti-GAPDH Primary | Abcam | cat# ab8245; RRID:AB_2107448 |
| Rabbit anti-WT1 Primary | Abcam | cat# ab89901; RRID:AB_2043201 |
| Rabbit Anti-mouse HRP Secondary | Abcam | cat# ab97046; RRID: AB_10680920 |
| Goat Anti-rabbit HRP Secondary | Abcam | cat# ab6721; RRID:AB_955447 |
| WT1 ChIP | Abcam | cat# ab89901; RRID:AB_2043201 |
| HA ChIP | Merck | cat# H6908; RRID:AB_260070 |
| H3K27Ac ChIP | Abcam | cat# ab4729; RRID:AB_2118291 |
| ETO (C terminus) ChIP | Santa Cruz | cat# sc9737; RRID:AB_2184227 |
| RUNX1 (C terminus) ChIP | Abcam | cat# ab23980; RRID:AB_2184205 |
| EGR-1 ChIP | Cell Signaling | cat# 4153; RRID:AB_2097038 |
| Sp1 ChIP | Sigma-Aldrich | cat# 17-601; RRID:AB_916354 |
| Critical commercial assays | | |
| Nucleobond spin kit | Machery-Nagel | cat# 740955 |
| RNeasy Plus micro kit | QIAGEN | cat# 74034 |
| TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero kit | Illumina | cat# 20020596 |
| NEBNext Ultra II RNA library kit | New England Biolabs | cat# E7770S |
| KAPA Hyper Prep kit | Roche | cat#KK8500 |
| QIAGEN MinElute Kit | QIAGEN | cat# 28604 |
| QIAGEN spin miniprep kit | QIAGEN | cat# 27104 |
| Nucleobond Xtra Midiprep kit | Machery-Nagel | cat# 740420.50 |
| Experimental models: cell lines | | |
| Kasumi-1 | DMSZ | Cat#ACC 220; RRID: CVCL_0589 |
| HEK293T | DMSZ | Cat#ACC 305; RRID: CVCL_0063 |
| Experimental models: organisms/ strains | | |
| NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ (NSG) mice | Jackson Laboratory | N/A |
| Recombinant DNA | | |
| pTRIPZ | Dharmacon | cat# RHS4744 |
| pENTR-GFP | In House | N/A |
| pCW57.1 | Addgene | plasmid #41393; RRID: Addgene_41393 |
| LentiCRISPRv2 | Addgene | plasmid #52961; RRID: Addgene_52961 |
| p300-dCas9 | Addgene | plasmid #83889; RRID: Addgene_83889 |
| tat, rev, gag/pol and vsv-g plasmids | Richard Mulligan (Mostoslavsky et al., 2005) | N/A |
| Software and algorithms | | |
| Bowtie | Langmead and Salzberg, 2012 | http://bowtie-bio.sourceforge.net/bowtie2/manual.shtml |
| MACS 2 | Zhang et al., 2008 | https://github.com/macs3-project/MACS |
| DEseq2 | Love et al., 2014 | https://bioconductor.org/packages/release/bioc/html/DESeq2.html |
| LIMMA | Ritchie et al., 2015 | http://bioconductor.org/packages/release/bioc/html/limma.html |
| HOMER | Heinz et al., 2010 | http://homer.ucsd.edu/homer/index.html |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| HiCUP | Wingett et al., 2015 | https://www.bioinformatics.babraham.ac.uk/projects/hicup/ |
| GOTHiC | Mifsud et al., 2017 | http://bioconductor.org/packages/release/bioc/html/GOTHiC.html |
| R scripting language | N/A | https://www.r-project.org/ |
| Picard | N/A | http://broadinstitute.github.io/picard/ |
| Bedtools | Quinlan, 2014 | https://bedtools.readthedocs.io/en/latest/ |
| Cytoscape | Shannon et al., 2003 | https://www.cytoscape.org/. |
| Trimmomatic | Bolger et al., 2014 | https://github.com/timflutre/trimmomatic |
| UCSC Genome Browser | Kent et al., 2002 | https://genome.ucsc.edu/ |
| STAR | Dobin et al., 2013 | https://github.com/alexdobin/STAR |
| Cufflinks | Trapnell et al., 2010 | http://cole-trapnell-lab.github.io/cufflinks/ |
| SUPPA2 | Trincado et al., 2018 | https://github.com/comprna/SUPPA |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for reagents and resources should be directed to, and will be fulfilled by the Lead Contact, Constanze Bonifer (c.bonifer@bham.ac.uk).

### Materials availability

The plasmids used in this study are available from the Lead Contact.

### Data and code availability

All high throughput data (RNA-seq, ChIP-seq, ATAC-sec and DNase I-seq data) generated in this study are available at the NCBI Gene Expression Omnibus (GEO) under the accession number GEO under accession number GSE153865. The published article includes AML patient RNA-seq and DNaseI-seq data from Assi et al., 2019b with GEO accession number: GSE108316.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Cell lines and cell line culture

Kasumi-1 and Human Embryonic Kidney (HEK) 293T cells were purchased from DMSZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen) and periodically checked for Mycoplasma infection and their identity checked by Short Tandem Repeat analysis. Cells were cultured in RPMI-1640 media (Kasumi-1) or DMEM (HEK293T) supplemented with 10% fetal calf serum (FCS) (GIBCO), 2 mM L-Glutamine and 100 U/ml Penicillin and 100 μg/ml Streptomycin. All cells were grown in a humidified incubator with 5% $CO_2$ at 37°C.

### Patient samples, purification and culture

AML and healthy patient blood or bone marrow were taken freshly from University Hospitals Birmingham with specific consent from the patients and ethical approvals from the National Health Service Research Ethics Committee.

Following acquisition of the patient sample, mononuclear cells were immediately isolated using density gradient centrifugation using Lymphoprep (StemCell Technologies). $CD34^+$ cells were purified using $CD34^+$ microbeads (Miltenyl-Biotech). Sorted $CD34^+$ cells were cultured in 90% StemSpan SFEM II (StemCell Technologies, 10% $CD34^+$ expansion supplement (StemCell Technologies) and 175nM UM171 (StemCell Technologies). Cells were grown in a humidified incubator with 5% $CO_2$ at 37°C.

## METHOD DETAILS

### Targeted myeloid sequencing panel

gDNA was extracted by phenol-chloroform extraction of DNA. 50ng of purified DNA was used for custom myeloid sequencing panel library preparation (Illumina) and sequenced on a MiSeq (Illumina) in paired ended mode with 150 cycles for each of Read 1 and Read 2 to achieve around 5000x coverage. Bioinformatics analysis was performed using the TruSeq Amplicon app in BaseSpace (Illumina). After demultiplexing, the software used a custom banded SmithWaterman aligner to align the reads against the human hg38 reference genome. The somatic variant caller then performed variant analysis for the specified regions. Mutations were called by annotation of filtered variants to databases such as the Catalogue for Somatic Mutations in Cancer (COSMIC) (Tate et al., 2019).

### RNA Sequencing

RNA was extracted using the Nucleobond spin kit (Machery-Nagel) or in the case of very low input RNA extractions of < 10,000 cells, the RNeasy Plus micro kit (QIAGEN) was used. Library preparation was performed using the TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero kit (Illumina) for cell lines or the NEBNext Ultra II RNA library kit (New England Biolabs) for primary cells. Then Illumina sequencing was performed on a NextSeq 550 (Illumina) run in paired-end mode for 75 cycles for each of read 1 and read 2.

Raw RNA-seq reads were processed with Trimmomatic v0.36 (Bolger et al., 2014) in order to trim sequencing adaptors and low quality bases. The processed reads were then aligned to the human genome (version hg38) using STAR (Dobin et al., 2013) with default settings. Separate density profiles for the positive and negative strand were generated using bedtools (Quinlan, 2014). Gene expression was measured as fragments per kilobase of transcript per million mapped reads (FPKM) values using Cufflinks (Trapnell et al., 2013) and differentially expressed genes were extracted using the cuffdif function in cufflinks and limma R package (Ritchie et al., 2015) and then were quantile normalized. Genes were considered as expressed with p value $\leq$ 0.01 and FPKM greater than 1 in at least one sample. Genes were said to be differentially expressed genes if there was at least 2-fold change in expression between any two conditions under test. Once the differentially gene expression lists were identified, the FPKM values from cufflinks were used for further analysis. The cuffquant function in cufflinks was used to generate cbx files, cuffquant was used for quantification with a GTF file. GTF files were downloaded from the UCSC genome browser. Files in cbx format were used as input into cuffdiff and cuffnorm functions.

Hierarchical clustering of RNA-Seq samples and replicates was done by first calculating the Pearson correlation value for each pair of samples. The resulting correlation matrix was then hierarchically clustered using complete linkage clustering of the Euclidean distances, and finally plotted as a heatmap in R.

Kyoto encyclopedia for genes and genomes (KEGG) pathway enrichment was performed using clueGO tools (Bindea et al., 2009). A gene was considered to be differentially expressed if it had a greater than 2-fold change between experimental conditions, and a Benjamini-Hochberg adjusted p value < 0.05. The size of the nodes reflects the number of genes within the term. The color of nodes reflects the enrichment significance of the terms. The network was laid out using Cytoscape (Shannon et al., 2003).

Splicing at the transcript isoform and differential splicing were identified using SUPPA2 pipeline (Trincado et al., 2018). SUPPA2 is a tool to study splicing across multiple conditions and to generate transcript events and local alternative splicing events from an annotation file in GTF format. Initially SUPPA2 was used to generate the ioi file which provides the set of all transcripts from that gene from which the transcript relative abundance is calculated. For the generation of PSI (Percentage Spliced In) values, SUPPA2 reads the ioi file generated and the expression file to calculate the PSI value for each of the events. Then SUPPA2 calculates the magnitude of splicing change (ΔPSI) for all transcripts and provide a p value as a measure for its significance across multiple biological conditions, using the different replicates per condition.

### DNase-I hypersensitive site mapping

DNase I digestions were performed as previously described (Bert et al., 2007). Briefly cells were permeabilized in DNase I resuspension buffer (60mM KCl, 10 mM Tris pH 7.4, 15 mM NaCl, 5 mM $MgCl_2$ and 300 mM sucrose) and then DNaseI diluted in dilution buffer (60mM KCl, 0.4% NP40, 15 mM NaCl, 5 mM $MgCl_2$, 10 mM Tris pH 7.4 and 2mM $CaCl_2$ was added and incubated at 22°C for exactly 3 minutes. The digestion was terminated by adding cell lysis buffer (300 mM Sodium Acetate, 10 mM EDTA pH 7.4, 1% SDS and 1 mg/ml proteinase K). DNA was then purified using phenol-chloroform extraction. Optimally digested DNA then underwent library preparation using the KAPA Hyper Prep kit (Roche) and size-selected for 200-300 bp fragments. Libraries were sequenced on a NextSeq 550 (Illumina) and run in single-end mode for 75 cycles.

Raw DNaseI-seq reads were trimmed with Trimmomatic v0.36 (Bolger et al., 2014) and aligned to the human genome version hg38 genome using Bowtie version 2.3.1 using default parameters (Langmead and Salzberg, 2012). Next duplicate reads and reads not uniquely aligned to chromosomal position were moved from aligned data using Picard v2.10.5. These filtered reads were used to generate density profiles using the genomeCoverageBed function from bedtools (Quinlan, 2014). These tag densities were displayed using the UCSC Genome Browser (Kent et al., 2002).

DHSs were called with MACS2 (Zhang et al., 2008) using the callpeak function (nomodel, call-summits and q = 0.005 parameters). The resulting peaks were then filtered against the hg38 blacklist and simple repeat tracks from the UCSC table browser (Karolchik et al., 2004) to remove any potential artifacts. Peaks were annotated to the nearest gene or using promoter Capture HiC data, where available, from (Assi et al., 2019b). The CHi-C paired-end sequencing reads from two AML patients and healthy $CD34^+$ cells were aligned using the *HiCUP* pipeline (Wingett et al., 2015) against the human genome (hg38). Statistically significant interactions were called using *GOTHiC* package (Mifsud et al., 2017) and HOMER software (Heinz et al., 2010). A merged set of significant interactions were identified and then were used to annotate ATAC and ChIP peaks to their genes.

Clustering of DNase I-seq samples was carried out using the merged set of peaks. The number of reads that mapped to these sites was counted in a 400bp window centered on the peak summit, and subsequently normalized to total sample size using DEseq2 (Love et al., 2014). Pearson correlation coefficients were then calculated between each pair of DNase I-Seq samples using the log2 of the normalized read counts, and then hierarchically clustered using Euclidean distance and complete linkage clustering of the correlation matrix in R. A peak was considered to be differentially accessible if the fold-difference of the normalized tag count was greater than 2 between experiments.

To create read density plots, peaks were first ordered according to fold-difference. The read density in a 2kb window centered on the peak summits was then calculated using from the bedGraph files produced by Homer using the annotatePeaks.pl file in HOMER, using the options -size 2000 -hist 10 -ghist. These were then plotted as heatmaps using java TreeView v1.1 (Saldanha, 2004).

De novo motif analysis was performed on peaks using HOMER (Heinz et al., 2010), recording the top enriched motifs within 200 bp from the peak summit. The annotatePeaks function in HOMER was used to find occurrences of motifs in peaks and known motif position weight matrices (PWM) were used from the HOMER database or from ChIP-seq experiments. PWMs of motifs (identified by HOMER) that are relatively enriched in a set of specific DNase I-seq peaks were used to calculate motif enrichment. For a given set $j$ of peaks, we defined a motif enrichment score ($S_{ij}$) for motif $i$ in peak set $j$ as

$$S_{ij} = \frac{n_{ij}/M_j}{\sum_j n_{ij} / \sum_j M_j}$$

where $n_{ij}$ is the number of peaks in each subset $j$ ($j = 1,2,\ldots$) containing motif $i$ ($i = 1, 2,\ldots,I$), $I$ is the total number of motifs used in the test, and $M_j$ the total number of peaks in each subset $j$ ($j = 1,2,\ldots$).

Motif co-localization clustering was performed as previously described (Ptasinska et al., 2014). A motif position search was done within WT1 ChIP-seq peaks. The distance between the centers of each motif pairs was calculated and the motif frequency was counted if the first motif was within 50bps distance from the second motif. Raw z-scores were calculated from the mean and standard deviation of motif frequencies observed in random sets using bootstrap analysis, peak sets with a population equal to that of the ChIP peaks were randomly obtained from the merged DNase I-seq peaks of all samples peak sets. Motif search and motif frequencies calculations were repeated 1000 times for each random set. A matrix was generated and Z-scores were displayed after hierarchical clustering as a heatmap with R.

### Assay for transposase accessible chromatin (ATAC)-seq

Omni ATAC-seq was performed as in Corces et al. (2017). Briefly, cells were washed in ATAC resuspension buffer (RSB) (10mM Tris-HCl pH7.5, 10mM NaCl and 3mM $MgCl_2$) and then lyzed for 3 minutes on ice in RSB buffer with 0.1% NP-40, 0.1% Tween-20. Then the cells were washed with 1ml of ATAC wash buffer consisting of RSB with 0.1% Tween-20. Then the nuclear pellet was resuspended in ATAC transposition buffer consisting of 25μl TD buffer and a concentration of Tn5 transposase enzyme (Illumina) related to the number of input cells, 16.5 μl PBS, 5 μl water, 0.1% tween-20 and 0.01% digitonin and then incubated on a thermomixer at 37°C for 30 minutes. The transposed DNA was then amplified by PCR amplification up to ½ of maximum amplification, as assessed by a qPCR side reaction. The library was purified and run on an Illumina Flow Cell in as per the DNase I sequencing protocol. Bioinformatic analysis of ATAC-seq was performed as per DNase I seq analysis with a few exceptions: 1) Bowtie2 was used with the parameter–very-sensitive-local. 2) Peak calling with MACS2 v2.1.1 used the settings–nomodel–nolambda -B–trackline.

### Chromatin immunoprecipitation (ChIP)-seq

Cells were crosslinked in 1% formaldehyde for 10 minutes before being quenched by 0.4 M Glycine. For EGR-1 ChIP, cells were also double-crosslinked with Di(N-succinimidyl) glutarate (DSG) for 45 minutes. Cells were lysed for 10 minutes in cell lysis buffer (5 mM PIPES pH 8.0, 85 mM KCl, 0.5% NP-40, 1:200 Protease Inhibitory Cocktail [PIC] before undergoing nuclear lysis for 10 minutes in nuclear lysis buffer (50 mM Tris pH 8.0, 10 mM EDTA, 1% SDS, 1:100 PIC). Nuclei were then sonicated in a Bioruptor Pico (Diagenode) until chromatin was sheared to a 200-600 bp fragment size and then diluted in blocking buffer (16.7 mM Tris pH 8.0, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton X-100, 0.01% SDS and 1:100 PIC). 15 ul of Dynabeads G were pre-blocked for 2 hours at 4°C with blocking buffer consisting of 0.5% BSA, 30 μg glycogen and ChIP grade antibody at a suitable concentration, before being incubated overnight 4°C with the diluted chromatin. See Key resources table for details of antibodies used.

Then the beads were washed in sonication wash buffer (50 mM Tris pH 8.0, 10 mM EDTA, 1% SDS), twice in Buffer A (20 mM Tris, pH 8.0, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl), twice in Buffer B (20 mM Tris, pH 8.0, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl), with Lithium buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA pH 8.0, 0.5% NP-40 and 0.5% sodium deoxycholate). Beads were then washed once in 1xTE. Chromatin was then eluted in 100mM $NaHCO_3$ and 1% SDS and reverse crosslinked for 4-16 hours at 65°C with 50 μg Proteinase K. DNA was then cleaned up with 1.8x SPRI beads. qPCR amplification of immunoprecipitated DNA was compared to input at several loci of putative protein binding sites as well as non-binding sites to assess the success of ChIP. See Data S4 for qPCR primer sequences.

Immunoprecipitated DNA was then library prepped with the KAPA Hyperprep kit as described before, except with a 200-450 bp size selection of the final library and Illumina sequenced as described before.

Reads from ChIP-Seq experiments were processed, aligned to the human genome and filtered in the same way as with DNase I-seq experiments. Peaks from transcription factor ChIP-seq experiments were identified using MACS2 with default settings. These peaks were then compared to the DNase I-seq or ATAC-Seq data, with only peaks that occurred within open chromatin regions being retained for further analysis. To identify differential binding of transcription factors such as WT1, initially union of all peaks was constructed by merging peaks that had summits within 100 bp of each other. The read density in these peaks was then retrieved using the

annoatePeaks.pl function in HOMER and normalized as counts per million in R. Peaks that had a fold-difference of at least 2 were considered to be differentially bound between experiments. The target genes for these peaks were identified by annotating each peak to its closest TSS using the annotatePeaks.pl function in Homer or using promoter Capture HiC data, where available, from Assi et al. (2019b) and Ptasinska et al. (2019). Motif enrichment analysis was performed as with DNase I-seq data.

Average profiles of ChIP-Seq data were constructed by first normalizing each of the alignment tracks as counts per million (CPM) using the bamCoverage function in deepTools v3.2.0 (Ramírez et al., 2014). These were then plotted using the plotProfile function in deepTools.

### Transcription factor and gene regulatory network identification

A subset of transcription factor (TF) genes and a gene regulatory network were identified as described in Assi et al., (2019b). The gene names for transcription factors in human were obtained from AnimalTFDB (Zhang et al., 2011). TFs and up or downregulated genes were considered as nodes and the nodes colored according to their gene expression fold-change. Node shape implies whether a gene is TF or a non TF gene. Node border color signifies whether the gene is either a RUNX1-ETO target or not. A directed edge from $TF_a$ to $TF_b$ or up/downregulated genes indicates motif binding of a $TF_a$ to the locus of the $TF_b$ and the edge is prominently displayed if $TF_a$ binds to the locus at that stage. The edge is classified and color coded according to the significant of motif count enrichment.

Motif count enrichment were defined as previously described in Assi et al. (2019b). Motifs were annotated to their related promoter using CHi-C data. Motif search within footprint coordinates where performed using HOMER (Heinz et al., 2010). The number of motifs per TF gene were counted and the significance of motif enrichment was identified using bootstrapping on random sampling, a random set of mapped motif were extracted from all union motif. After 1000 iterations the mean, standard deviation and the z-scores are computed. Motif ($TF_a$) is linked to gene or $TF_b$ with positive z-score values only.

### siRNA knockdown

Cells were resuspended to a concentration of $15x10^6$ cells/ml and transferred to a 4mm electroporation cuvette. 200nM of siRNA was added to the cuvette and then cells underwent electroporation for 10 ms at 350V in a square wave electroporator. Electroporation was repeated every 2 days. See Supplementary file 5 for a table of siRNA sequences.

### Cloning of shRNA constructs

Custom linker A and linker B oligonucleotides were designed as in Dow et al. (2012) with sense and anti-sense sequences. See Supplementary file 5 for a table of shRNA sense sequences. 100 pmoles of linker A and linker B were phosphorylated and annealed using T4 Polynucleotide Kinase and a gradual cooling ramp from 95°C to 25°C. pTRIPZ (Dharmacon) was digested with XhoI and MluI restriction enzymes and then dephosphorylated with Calf Intestinal Alkaline Phosphatase. The oligonucleotides digested plasmid were then ligated using T4 DNA Ligase. The annealed plasmid was then transformed into competent bacteria, grown in Lysogeny Broth and underwent maxiprep using the Nucleobond Xtra Midiprep kit.

### Cloning of overexpression constructs

WT1 isoform cDNA was a gift from Naohiko Seki, Chiba University, Japan (Moriya et al., 2008). Dominant negative FOS (dnFOS) cDNA was a gift from Charles Vinson, National Cancer Institute, Bethesda, USA (Olive et al., 1997). The cDNA was amplified by PCR using primers with SalI and NotI restriction enzyme overhangs and then digested with SalI-HF and NotI-HF. See Supplementary file 5 for a table of PCR primers.

pENTR-GFP plasmid was digested with SalI and NotI and the larger 3925 bp fragment was gel extracted and then ligated to the dnFOS or WT1 isoform using T4 DNA ligase. The ligated plasmids were then recombined in the tet-on plasmid pCW57.1 (David Root, Addgene plasmid #41393) using Gateway technology. The recombined plasmid was then transformed into competent bacteria, grown in Lysogeny Broth and underwent maxiprep using the Nucleobond Xtra Midiprep kit.

### WT1 mutant cloning

*WT1* mutant expressing plasmids were created with stop codons in exons 7, 8 or 9 by site-directed mutagenesis of pENTR-GFP-WT1 by PCR using primers. See Supplementary file 5 for a table of PCR primers. Then the plasmid was recombined into pCW57.1 and amplified and maxiprep performed as before.

### HA tagged WT1 cloning

A HA tag was added to either the N-terminal end or C-terminal end of WT1 by site-directed mutagenesis of pENTR-GFP-WT1 by PCR using primers. See Supplementary file 5 for a table of PCR primers. Then the plasmid was recombined into pCW57.1 and amplified and maxiprep performed as before.

### CRISPR activation

LentiCRISPRv2 plasmid (Feng Zhang, Addgene plasmid #52961) and p300-dCas9 (Charles Gersbach, Addgene #83889) were digested using RsrII and AgeI and the 6241bp band was gel extracted from p300-dCas9 and the 10495bp band from LentiCRISPRv2. These two fragments were then ligated using T4 DNA ligase to give LentiCRISP-p300-Cas9, transformed into chemically competent bacteria and amplified and maxiprep performed as before.

Linker A and linker B oligonucleotides of the following sequences but with BsmBI overhangs were annealed and ligated into LentiCRISPR-p300-Cas9 using T4 DNA ligase, transformed into chemically competent bacteria and amplified and maxiprep. See Supplementary file 5 for a table of CRISPR crRNA sequences.

Following lentiviral transduction of Kasumi-1 cells, cells were assessed for H3K27Ac deposition by the dCas9-p300 by ChIP-qPCR. Enrichment was normalized to a known site of H3K27Ac deposition, the PU1 −14 kb enhancer and to enrichment of the empty vector control. See Supplementary file 5 for a table of ChIP qPCR primer sequences.

### Lentiviral transduction

tat, rev, gag/pol and vsv-g plasmids were a gift from Richard Mulligan, Harvard medical school, USA (Mostoslavsky et al., 2005). HEK293T cells were grown in adherent culture and transfected using calcium phosphate co-precipitation: five plasmids (pCW57.1 with cDNA insert, tat, rev, gag/pol and vsv-g) were mixed with 250 mM Calcium Chloride and 2.5 mM HEPES. This was then slowly added to a buffer consisting of 50 mM HEPES, 280 mM Sodium Chloride and 1.5 mM $Na_2HPO_4$ and incubated at room temperature for 35 minutes and then promptly added dropwise to the plates of HEK293T cells. Viral supernatant was then harvested after 24 hours and subsequently every 12 hours for 36 hours prior to concentration. Virus was concentrated using a Centricon Plus-70 100kDa filter.

$1x10^6$ cells from a cell line or sorted $CD34^+$ primary cells were mixed with concentrated viral supernatant and 8 μg/ml polybrene before being spinoculated at 1500xg at 32°C for 90 minutes. If the viral constructs contain a puromycin resistance cassette, cells were then puromycin selected for 7 days or alternatively cells transduced with a lentivirus with a GFP cassette were sorted by FACS on a FACS Aria Custom machine.

### Colony formation assays

In the case of cell lines, 2000 cells were seeded into 1.1 mL of a pre-mix of 2.6% Methocult methylcellulose (StemCell Technologies), 20% fetal calf serum, 0.4 mM L-Glutamine, 25 mM HEPES and 40% Iscove's Modified Dulbecco Medium with or without 2 μg/ml of Doxycycline. Colonies were assessed and counted 10 days after incubation at 37°C. In the case of primary cells, 2000 cells were seeded into 1.1ml of Methocult Complete with or without 2 μg/ml of Doxycycline. Colonies were assessed and counted 14 days after incubation at 37°C.

### Growth curves of cell lines

$2.5x10^5$ cells/ml were cultured in usual culture media. Cells were mixed with an equal volume of 0.4% trypan blue. Live cells which did not uptake trypan blue were counted and cells were split every 2 days to $2.5x10^5$ cells/ml and 2 μg/ml of doxycycline was added where appropriate.

### Flow cytometry

All flow cytometry was performed on a CyAn ADP machine (Beckman Coulter) and analyzed on Summit 4.3 (Beckman Coulter). Cell differentiation was assessed by antibody staining for the stem cell marker CD34-PE (Miltenyi Biotech, cat# 130-113-179) and the differentiated myeloid cell marker CD11b-APC (Miltenyi Biotech, cat# 130-110-554). Apoptosis was assessed by using Annexin V kits with Annexin V-FITC (BD Biosciences, cat# 556547) or Annexin V-APC (BD Biosciences, cat# 550474) in experiments in any cells were already GFP positive. Cell cycle analysis to distinguish the G1, S and G2-M phases of the cell cycle was performed by analyzing cellular DNA content as assessed by fluorescence from the DNA intercalating stain 7-aminoactinomycin D (7-AAD).

### Cytospins

5000 cells were spun onto a slide at 800 rpm for 4 minutes using a cytospin 4 cytocentrifuge (Thermo Scientific). Cells were then methanol fixed and stained with Wright-Giemsa stain.

### Western blot

Protein extraction was carried out by lyzing cells in RIPA (Radioimmunoprecipitation assay) buffer. Protein was mixed with Laemmli buffer and was denatured at 95°C. The protein was then separated a 5%–20% Tris-Glycine-Extended gel in SDS-PAGE running buffer (0.025 M Tris pH 8.3, 0.192 M glycine and 0.1% SDS). The protein was then transferred onto a nitrocellulose membrane which was blocked with 10% milk before being incubated with a primary antibody diluted in 5% milk at 4°C for at least 16 hours. See key resources table for antibodies used. Following this, the membrane was incubated with the appropriate Horse Radish Peroxidase (HRP) conjugated secondary antibody in 5% milk and the membrane developed using the Pierce Enhanced Chemiluminescence kit (Thermo Scientific) and imaged on a Bio-Rad ChemiDoc XRS+.

## QUANTIFICATION AND STATISTICAL ANALYSIS

If not indicated otherwise, all statistical comparisons were performed using two-sided Student's t test.

Differential gene expression and differential chromatin accessibility analysis was carried out using Limma. A gene was considered to be differentially expressed or differentially accesible if it had a greater than 2-fold change between experimental conditions, and a Benjamini-Hochberg adjusted p value < 0.01.